

# **CHAPTER 1**

## **GENERAL INTRODUCTION**

## ***1.1. Introduction***

Commercial beef production in Argentina is mainly based on cow-calf production and/or fattening systems, in both temperate and subtropical environments.

Breeding objectives for these systems should be set by seed livestock breeders according to the conditions and personal desires of their bull-buyers. These objectives could be affected by soil type, environment, genetic merit of the female herd, operational and financial constraints, target beef markets, etc.

According to these breeding objectives, a breeding program is designed, by which the most suitable individuals in the current generation are selected and mated to obtain a genetically improved next generation. So, after establishing these breeding objectives, individual mating allocations need to be made to efficiently meet them. A Mate Selection program establishes which animals to select as parents and which sire should be mated to which dam according to an objective function to be maximized. This function takes into account the availability and costs of reproductive mating technologies (natural mating, MOET, AI, etc.) and, furthermore, factors like progeny merit, inbreeding and genetic diversity levels, production costs, levels of investment, etc. The mating strategy could also lead to divergent lines serving different segments in the client market.

In this way, Mate Selection becomes a complex task, even more as the number of variables involved in the breeding program increases (number and size of herds, number of traits considered, type of costs, etc.). To solve this, special techniques or algorithms suitable for global numerical optimisation are used. One of the most efficient of these algorithms is Differential Evolution.

Differential Evolution is related to Evolutionary Computation which comprises algorithms that emulate the biological processes of evolution. This means that they use mechanisms analogous to selection, mutation and recombination (crossing over) in repeated cycles or generations.

In order to implement and optimize cattle breeding program design two types of approaches have been developed. The first type of these approaches has been referred to as tactical, which are tools for animal breeding concerned with day-to-day, or operational, breeding program decision-making. On the other hand, the second type of approaches is often referred to as strategic, which aim to predict response to selection under different breeding program designs and evaluate different selection and recording strategies. Strategic methods are used before the breeding program is started, and intermittently to review new developments, rather than on a day-to-day basis.

This study will focus on a comparison of results of applying tactical and strategic approaches for different breeding programs, using a Differential Evolution algorithm for Mate Selection, according to diverse objectives targeting commercial beef production systems in Argentina.

Different outcomes of improvements of these systems will be defined. Examples of outcomes to be targeted are: overall merit, individual trait merit, genetic diversity to target multiple beef markets, and variation in bulls produced (e.g. variation in overall merit, or across traits, or pedigree), etc.

It should be noted that although the cases presented herein involve beef cattle, the approaches discussed within this study are broadly applicable across species.

## ***1.2. Outline of thesis***

The purpose of the study is to compare results of applying tactical and strategic approaches for different breeding programs with diverse implementations which target possible commercial production scenarios in Argentina, using a Differential Evolution algorithm for Mate Selection in a pure beef cattle population.

Animal breeding aims to give genetic improvement of livestock populations in order to produce the desired animal products in an efficient and

sustainable way considering the present and expected future conditions of the target markets. To achieve this goal it is necessary to select the most suitable individuals in the current generation and mate them in such a way that a new generation is obtained in which is observed a genetic response in the desired direction or directions.

Following this, the main concepts of breeding program design are described in Chapter 2, reviewing the different steps in order to set a suitable breeding program for beef cattle production and the methods for the evaluation of the breeding program design. At the same time, diverse methods for the implementation of breeding objectives are discussed, reviewing the concepts related to setting suitable selection indices to help achieve the desired breeding objectives.

Mate Selection is a complex task, even more so as the number of variables involved in the breeding program increases (number and size of herds, number of traits considered, type of costs, etc.). At the end of the Chapter 2, global numerical optimisation related to evolutionary algorithms is discussed in relation to application in Mate Selection.

Chapter 3 presents the results of some simple examples of Mate Selection using the Pedigree Viewer system (<http://www-personal.une.edu.au/~bkinghor/pedigree.htm>) based on a simulated population of beef data. This shows the impact of different policies, constraints and parameter settings. Then, using the same simulated beef data, the results of aiming to improve the immediate next generation are compared with a simple look-ahead scheme that also aims at the grand progeny generation by maximizing the genetic selection differential to be achieved in the progeny generation (following Kinghorn and Shepherd, 1994 and Hayes et al., 1998). This will set the scene for investigating the setting up of a Young Sire Program using Look-Ahead Mate Selection (LAMS) later in Chapters 5 and 6.

With the interest to explore diverse aspects of Mate Selection in practical situations, in Chapters 4, 5 and 6 three experiments are carried out, reflecting typical beef production scenarios.

Herd size is a factor that affects genetic response due to its relation to selection intensity and inbreeding, especially in the long term, through effects on effective population size. Also the availability of more phenotypes probably leads to more accurate genetic evaluation of animals. According to this, two types of breeding enterprises of purebred beef cattle exist in Argentina. One corresponds to stud-farms with a moderate to large herd size that apply a two-tier nucleus breeding scheme, in which the nucleus is closed to the lower tier but open to import animals, semen or embryos from other nuclei. In Chapter 4 the consequences of applying a tactical mate selection approach in this scenario using real data from leader Angus purebred stud-farms of Argentina are analysed, considering different kinds of costs and constraints.

It should be taken into account that all the cases in Chapters 3 and 4 are analysed after only one round of selection.

On the other hand, in Argentina there exists a considerable proportion of stud-farms with a limited herd size. For these farms a cooperative Young Sire Program (YSP) is proposed in order to improve the rate of genetic gain without increasing inbreeding. In Chapter 5 different aspects of the implementation of this YSP are analysed through a deterministic modelling and simulation. In Chapter 6 the consequences of applying a look-ahead mate selection scheme approach in this scenario are tested stochastically. In these two last mentioned Chapters simulations are done over multiple selection rounds, being the approach fully tactical as the solution is optimised at every selection round. The key question here will be whether tactical Mate Selection can in fact set up such a scheme, and whether it improves on outcomes by making fully opportunistic use of the prevailing animal resources. The idea of setting up a YSP is not only due to its advantages, being possibly one of them that using a higher proportion

of young sires might be cheaper, but also because it accommodates to the LAMS approach, by which extra genetic gains might be obtained.

It should be noted that the LAMS approaches used in Chapter 3, on one hand, and Chapters 5 and 6, on the other, differ. In Chapter 3, LAMS aims for the maximization of the selection differential achieved among the progeny, which reflects genetic merit in the grand progeny generation. This is predicted by the difference between the average merit of their selected parents and the average merit of the entire population of current candidates before they are selected. To do this, the progeny predicted to be born from the current selected candidates are themselves put through a selection process, and the means of these selected progeny used to predict mean grand progeny merit. On the other hand, Chapters 5 and 6 aim for the maximization of the grand progeny mean considering both mean EBV of the selected parents and the predicted residual genetic variance within each progeny family, as impacted by both selection accuracy (and inbreeding for Chapter 6) in the selected parents.

Finally, the general discussion in Chapter 7 reviews the use of genetic algorithms as a tool for Mate Selection decisions and the optimisation of breeding program designs of a wide range of complexity.

## **CHAPTER 2**

### **LITERATURE REVIEW**

## ***2.1. The breeding program design***

The purpose of animal breeding is the genetic improvement of livestock populations in order to produce the desired animal products in an efficient and sustainable way taking into account the present and expected future conditions of the market. This is achieved by selecting the most suitable individuals in the current generation and mating them in such a way to give a sustainable genetic response in the desired direction.

A breeding program is the organized structure that is set up in order to realize the desired genetic improvement of the population.

A breeding program has the following components:

- a) A data recording system on selection candidates to identify the best individuals,
- b) Methods and tools for breeding value estimation of selection candidates, i.e. a genetic evaluation system,
- c) A system to select the animals that become parents of the next generation, and mate them to produce the next generation, and
- d) A structure to disseminate the genetic improvement created in the breeding population into the production population.

### **2.1.1. Breeding objective**

As previously mentioned, the aim of a breeding program is to select the most suitable individuals in the current generation and use them as parents to obtain a genetically improved next generation.

“Suitable” or “best” are relative terms, which mean that depending on the environmental conditions, genetic merit of the female herd, management practices, operational and financial constraints and target markets, animals with

different genotypes can be classified accordingly. In addition, the personal desires of the bull-buyers should be taken into account as they can largely dictate the kind of animal to produce. Breeders at each level of the breeding structure (i.e. elite breeders, multipliers and commercial producers) try to produce animals that will be in the greatest demand by their customers at the next level down, with the ultimate result that the best animal is the animal that is the most useful or profitable for the end user. End users can thus be defined as the individuals whose particular needs should form the basis for determining breeding objectives.

In the beef industry, the end users are commercial producers. These are the people whose primary products are commodities for public consumption. Commercial producers are in most cases not the end of the production chain; beyond them are the processors (slaughterhouses), the retailers and consumers (domestic and external markets). But the commercial producers are end users because their particular needs reflect the requirements of the entire production chain. They need animals that are physically and reproductively sound, healthy and perform efficiently in their environment. They also need animals that possess the product and performance characteristics required by the retailers and consumers. The importance of these latter characteristics should be reflected in the prices paid to the commercial producers for their products. In the Western world, the interest of consumers in the system of production has increased over time. This increased awareness of consumers has resulted in an increased emphasis on health and welfare traits in the breeding objective of farm animals and reduced emphasis on primary production traits (e.g. growth rate).

So, the breeding objective can be defined as the overall goal of the genetic improvement program to maximize profit, or to maximize economic efficiency or to minimize economic risk (Cunningham, 1979; Ponzoni, 1982; Harris et al., 1984).

Genetic change takes place through selection. Selection criteria can be derived from information on animals' phenotypes and those of their relatives.

Phenotypic information is generally based on objective measurement of animals' characteristics, but it can also include subjective scoring. Even genotype information about DNA-markers can be put in this category as it contributes to predicting an animal's breeding value.

In most cases more than one trait that will contribute to the overall breeding objective is to be improved. So, having defined the breeding objective, it becomes necessary to define the relative importance of these traits (which are called selection criterion traits). This is done by determining the economic weight for each of those traits (Hazel, 1943). For a given animal that is a candidate for selection, the sum of its additive genetic value multiplied by the economic weight for each trait is referred to as the aggregate genotype. In this way, the aggregate genotype is a mathematical function of genetically controlled traits (i.e. the selection criterion traits) that when use appropriately in a breeding program will maximize the breeding objective.

Thus

$$H = v_1g_1 + v_2g_2 + v_3g_3 \dots \dots + v_mg_m$$

where  $H$  is the aggregate genotype for an animal,  $v_m$  are economic weights of  $m$  traits and  $g_m$  are the animal's breeding values of  $m$  traits for all those traits included. The economic weight of a given trait is defined as the rate of change in profit per unit improvement in that trait when holding all other traits unchanged.

The economic weight of each trait can be estimated directly from the production model, by calculation of partial first derivatives of the profit function or applying bio-economic models in which relevant biological and economic aspects of the production system (Nitter et al., 1994; Cunningham et al., 1970; Tallis, 1962; Kempthorne and Nordskog, 1959; Koots and Gibson, 1998; Yamada et al., 1975; Fowler et al., 1976; Sivarajasingham, 1995) are modelled as a system of equations.

In the process of choosing the traits that will be included in the aggregate genotype all traits that directly contribute to the breeding objective should be considered, while those that have an indirect impact on the objective (e.g. indicator traits) should not be included – their impact on the breeding program can be mediated as measured selection criterion traits. In addition, traits that have little or no genetic variation do not need to be taken into account.

### **2.1.2. Implementation of breeding objectives**

Breeding objectives indicate the traits to be improved and the direction for genetic change.

As was mentioned above, with the traits of commercial importance included in the breeding objective (i.e. the selection criterion traits) are allocated economic weights that describe the economic impact of a unit change in each trait. These economic weights need to be expressed in the same economic units, such as dollars profit per breeding cow per year.

There are two approaches to calculate the economic weights:

- a) Economically Rational approach, and
- b) Desired Gains approach.

#### **2.1.2.1. The Economically Rational approach**

A strategic implementation of a breeding objective derives from an economic approach to the latter. So certain rules and steps are followed to achieve the breeding goal.

In order to define the breeding objective through an economic approach it is necessary to take into account the costs and returns of the animal production process so the economic importance of different traits can be defined. In addition the future possible production conditions and variation of the

costs/returns relationships should be considered. This increases the uncertainty of any definition of breeding objectives, which is especially true in countries with a strong or even dominating influence of the government on the economy.

It should be chosen to include in the breeding objective all those traits with a major impact on the efficiency of commercial production. This must be done only based on economic criteria. The genetic parameters are considered later when the breeding values are estimated. It does not matter whether the traits chosen can be measured or whether auxiliary traits will be utilized to predict them whenever there are genetic correlations to those traits. This can also prevent the occurrence of unwanted detrimental effects due to genetic antagonisms.

As the economically rational approach does not consider genetic parameters to calculate economic weights, the value of changing one unit of a trait is not influenced by how difficult it is to generate that change. So, the selection criteria are affected by genetic parameters, but not the economic weights themselves. In this setting, breeding objectives should reflect the costs and returns involved in a production system, and should not consider costs and gains generated in a breeding program (Graser, 2006).

The economic weight of each trait can be estimated directly from the production model, by calculation partial first derivatives of our profit function or applying bio-economic models in which relevant biological and economic aspects of the production system are described as a system of equations.

As the classical economically rational calculation of economic weights ignores genetic parameters, the resulting direction of genetic gain is not fully addressed. Moreover, the economically rational approach usually assumes linearity, which may not be the case, usually ignores the reality of the seed stock marketplace, and can have difficulty in handling some traits such as disease resistance (Kinghorn, 2005).

The economically rational approach is ideal where there is a strong knowledge about product value in the marketplace.

### **2.1.2.2. The Desired Gains approach**

The 'Desired Gains' approach is an alternative approach to developing breeding objectives. It involves declaration of the relative magnitudes of genetic gain desired in the traits of importance. The breeding objective calculations still result in relative weights, but these are now influenced by genetic parameters, with generally greater weightings for traits which are more difficult to change.

A convenient way of doing this is to view the range of possible selection response outcomes. This can be achieved by looping through combinations of economic weights for traits and plotting the predicted response in each trait to get a different ellipse per each pair of traits. Changing genetic parameters changes the shape of the ellipses. The selection index will continue to do the best job possible, but the direction of desired gain might change according to the different possible outcomes. The Desire program (Kinghorn, 2000), which is largely based on the method of Brascamp (1984), shows this feature properly. Linear programming can also be used to achieve constrained and desired gain indexes (Keller and Gibson, 1990; Toro, 1992).

The restricted index is a simple subset of the Desired Gains approach, in which the objective is set up to give a predicted no genetic change in one or more nominated traits. An example in beef cattle is restriction for no change in back fat thickness.

The Desired Gains approach takes into account genetic parameters in calculation of economic weights so direction of genetic gain can be fully addressed and accommodates practitioner feelings that the best direction to take depending on how far s/he can go in each direction. On the other hand, Desired Gains can be disturbed by irrational and dogmatic opinions (Kinghorn, 2005).

- **Tactical implementation of the desired gains approach**

There is a different approach, which makes predictions for the next progeny crop based not on genetic parameters, but on the prevailing list of estimated breeding values of candidate animals for selection. Predicted progeny merit is simply the average of parental EBVs, and there is no need to involve estimated genetic parameters. The Progeny Explorer program (Kinghorn et al., 2001) shows this feature properly (Kinghorn, 2005).

### **2.1.2.3. Selection Index**

In general, selection programs in livestock production aim for the simultaneous improvement of several traits. The theory to optimise selection on multiple traits is based on the selection index principle. The selection index approach to developing optimal selection criteria for profitability was first introduced by Hazel (1943).

Based on multiple trait selection index theory, information on phenotypes and pedigree or EBVs as well as economic weights is combined into one index value (the selection index), that is the best estimate of the aggregate genotype which will achieve the breeding objective.

While breeding value estimation is obtained by BLUP (Henderson, 1973), the selection index is used mainly to derive weights, to understand the relative value of different sources of information, and to predict response and selection accuracy for different alternatives. Thus, for example, selection index theory is used to calculate the additional merit of measuring a particular trait.

In multiple trait selection the relative economic importance of the different traits (i.e. the selection index traits) is reflected by economic weights, but the weights to be used in the selection index must also account for the genetic parameters that affect how much response can be made. The index is

then the best ranking for genetic merit for an aggregate genotype for the different traits in the breeding objective.

The index  $I$  is written as

$$I = b_1X_1 + b_2X_2 + \dots + b_nX_n$$

where  $X_i$  is a measurement of the phenotypic performance of an animal (or its relatives) on the  $i^{\text{th}}$  trait and  $b_i$  is the selection index coefficient (or weight) for that trait calculated by multiple regression:

$$b = P^{-1}Gv$$

where  $b$  is a column vector of  $m$  selection index coefficients or weights,  $P^{-1}$  is an  $m \times m$  inverse matrix of phenotypic covariances among the observations in the selection index,  $G$  is an  $m \times n$  matrix of genetic covariances among the  $m$  index observations and the  $n$  traits in the aggregate genotype and  $v$  is a column vector of economic weights or weighting factors of the  $n$  traits in the aggregate genotype.

The selection index weights ( $b$ 's) are calculated such that selection of individuals on their selection index value,  $I$ , maximizes response in the aggregate genotype,  $H$ , i.e. the correlation ( $r_{IH}$ , see later) between animals' selection index values and their dollar Breeding Values is maximized according to the prevailing objective.

It is necessary to take into account that each one of the considered selection index traits must be one of the chosen traits in the aggregate genotype (i.e. the selection criterion traits) or be genetically correlated to one or more traits in the aggregate genotype.

So, using for example two traits

$$I = b_1X_1 + b_2X_2$$

being

$$P = \text{var} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \text{var}(X_1) & \text{cov}(X_1, X_2) \\ \text{cov}(X_2, X_1) & \text{var}(X_2) \end{pmatrix} = \begin{pmatrix} \sigma_{P_1}^2 & r_P \sigma_{P_1} \sigma_{P_2} \\ r_P \sigma_{P_1} \sigma_{P_2} & \sigma_{P_2}^2 \end{pmatrix}$$

where  $\text{var}(X_i)$  is the variance of phenotypic observations for trait  $i$ , which is equal to the phenotypic variance for trait  $i$  (indicated as  $\sigma_{P_i}^2$ ). Likewise,  $\text{cov}(X_x, X_y)$  is the covariance between animals' phenotypic performances on these two traits, which is equal to the phenotypic covariance,

$$G = \text{cov} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, (g_1 \quad g_2) = \begin{pmatrix} \text{cov}(X_1, g_1) & \text{cov}(X_1, g_2) \\ \text{cov}(X_2, g_1) & \text{cov}(X_2, g_2) \end{pmatrix} = \begin{pmatrix} \sigma_{g_1}^2 & r_g \sigma_{g_1} \sigma_{g_2} \\ r_g \sigma_{g_1} \sigma_{g_2} & \sigma_{g_2}^2 \end{pmatrix}$$

and  $G$  is the covariance between phenotypic observations and breeding values and contains only genetic components: either the genetic variance (if same trait) or the genetic covariance (if different traits).

Then

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = P^{-1}Gv$$

where  $v$  is the vector of economic weights derived in the economically rational approach or the equivalent weighting factors of traits derived in a desired gains approach.

The same principle can be applied using the estimated breeding values (EBV) coming from a genetic evaluation program:

$$I = v_1 \text{EBV}_1 + v_2 \text{EBV}_2 + \dots + v_n \text{EBV}_n = v' \hat{u}$$

where  $v$  is a vector of economic weights or weighting factors and  $\hat{u}$  is a vector of estimated breeding values. Notice that in this case the economic weights or weighting factors can be used directly in the selection index equation because the EBVs have already been calculated with due attention to the phenotypic and genetic parameters.

It should be noted that the advantage here is that EBVs are usually more accurate than individual's phenotypes, as they have been corrected for fixed effects simultaneously with use of information from relatives.

The index value can be calculated for each animal and then they can be ranked on index. It would be expected that the highest ranked sire on index would be the most profitable, the second highest the second most profitable and so on.

Recall that selecting more sires will result in lower inbreeding, but selecting fewer sires results in higher selection intensity. Selecting more sires also spreads the risk of making errors in selection, with a more predictable outcome.

The criteria for inclusion of traits in the selection index are:

- The trait must be recorded such that EBV can be obtained on selection candidates
- The trait must have reasonable heritability, (although low heritable traits can provide accurate EBV if sufficient data are available, such as a progeny test).
- The trait must be one of the traits in the aggregate genotype or be genetically correlated to one or more traits in the aggregate genotype.

- **Accuracy of index selection**

The accuracy of selection is the correlation between the overall selection criterion (i.e. the selection index) and the true breeding value for the breeding goal that is to be improved.

Then, if the variance of the index is:

$$\text{var}(I) = \sigma_I^2 = \text{var}(b'X) = b' \text{var}(X) b = b'Pb$$

and the variance of the breeding objective is:

$$\text{var}(H) = \sigma_H^2 = \text{var}(v'g) = v' \text{var}(g) v = v'Gv$$

then, the covariance between the index and the breeding goal is:

$$\text{cov}(I,H) = \text{cov}(b'X,v'g) = b' \text{cov}(X,g) v = b'Gv = b'Pb$$

So, the covariance between the index and the true genotype is equal to the variance of the index. The correlation is then:

$$r_{IH} = \text{cov}(I,H)/\sigma_I\sigma_H = \sigma_I^2/\sigma_I\sigma_H = \sigma_I/\sigma_H = \text{SD\_Index} / \text{SD\_Breed\_Objective}$$

where  $r_{IH}$  is termed selection accuracy.

With high accuracies, the standard deviation of the index is almost equal to the standard deviation of the breeding goal, and with low accuracy the ratio is relatively lower.

Accuracy of selection explains genetic variation. Thus, low accuracy in a group of animals means a higher residual variance of true genetic merit that can be exploited as more information becomes available. This concept will be used later in this study (mainly in Chapters 5 and 6).

• **Response to index selection**

To be able to predict response to selection using the selection index it is necessary to know the variance of the index,  $var(I)$ :

$$var(I) = \sigma_I^2 = \mathbf{b}'\mathbf{P}\mathbf{b}$$

If selection is based on an index, the response (mean of selected group) will depend on selection intensity and the standard deviation of the index values.

$$\text{So, response} = R_{\text{gen}} = i\sqrt{\mathbf{b}'\mathbf{P}\mathbf{b}} = i\sigma_I = i r_{IH} \sigma_H$$

which is the response ‘per generation’ and ‘per round of selection’.

Also,

$$\text{the Response for trait } i \text{ becomes: } \quad ib'_i G_i / \sigma_i$$

The index and index accuracy can be different for males and females, as there may be more information measured on one of the sexes.

The Response per year (in dollars):

$$R = \frac{i_m \sigma_{Im} + i_f \sigma_{If}}{L_m + L_f} = \frac{i_m r_{IH_m} + i_f r_{IH_f}}{L_m + L_f} \sigma_H$$

The Response per year (for each trait):

$$R = \frac{i_m b'_m G_{mi} / \sigma_{lm}}{L_m + L_f} + \frac{i_f b'_f G_{fi} / \sigma_{lf}}{L_m + L_f}$$

where  $m$  and  $f$  refer to male and female index, and  $G_{mi}$  is the  $i^{th}$  column of the  $G$  matrix for the male selection index.

The units of response are dollars, as the index is expressed in dollars.

### ● Non-linear Profit and Selection

The use of economic values in a linear index assumes that profit is linearly associated with the traits.

Several studies have pointed to the weakness of assuming a linear profit function (Ponzoni, 1982; Bright, 1991; Amer et al., 1994).

There are many types of non-linear relationships. Some traits have an economic optimum, for instance carcass fat depth (Shultz, 1990; Hovenier et al., 1993).

Non-linear relationships among traits may cause a re-ranking of animals on index or profit.

The linearity of the breeding goal implies that a negative value for one trait can be compensated by a positive value for another trait in the breeding goal. From the perspective of an individual breeding company, which has to consider the competitive value of his livestock, this compensation does not always hold. If the animals are low for a certain trait, it may be under a certain acceptable level for the consumers. In that case, the improvement should be largely focused on this trait that deteriorates the competitive value, until it is above an acceptable level.

If the degree of nonlinearity is high, suboptimal profit may result.

Non linearity of economic value is for many breeders a reason to make recourse to other selection methods. A simple alternative is independent culling, a method that selects animals only if they are above a minimum value for a certain set of traits. However, for small deviation of linearity, selection index is still an optimal method, and independent culling is in most cases not as optimal as selection index (van der Werf, 2007). It should be noted that, in the face of such non-linear economics, breeders can always use desired gains to target give outcomes of given time horizons.

## ● **BLUP**

Best Linear Unbiased Prediction (BLUP) is a method developed by C.R. Henderson (1959, 1973, 1976, 1984, 1990) by which breeding values for commercially important traits are estimated (EBVs).

The BLUP procedure is based on the principles of selection index and linear models. Through selection index (also termed sometimes as Best Linear Prediction, or BLP), BLUP estimates breeding values by regression using phenotypic information on animals and their relatives, while linear models predict breeding values based on phenotypic information, correcting them for fixed effects. Thus, a mixed model is the base of BLUP estimation of breeding values, which is a linear model containing fixed effects as well as random effects (i.e. the additive genetic values). So, both effects (fixed and random) are fitted in the same model and estimated simultaneously in the same analysis.

For prediction of breeding values, BLUP uses all available information, typically including information from all known and recorded relatives. In the case that a multiple trait analysis is done, BLUP also uses information from correlated traits. In addition, it corrects EBVs for systematic environmental effects (called fixed effects) such as herd, year-season of production etc., and accounts for unequal use of the best sires in different herds, for selection and

non-random mating. BLUP requires the genetic parameters (heritability, correlations) of the traits to be evaluated for use as priors, and a good data structure. So, BLUP encompasses the statistical properties of the EBVs obtained using this method.

We can also mention as properties of BLUP that:

- it corrects for the breeding value of mates, e.g. for assortative mating,
- it accounts for possible genetic differences between identified fixed effect groups, taking deviations from expected genetic means, and regresses the EBV towards the expected genetic mean,
- it allows for estimation of the genetic trend, which is the mean EBVs per year of birth,
- it allows selection across age classes; satisfying the need to evaluate all animals jointly.

Because all this, the accuracy of EBVs is maximized, but also BLUP accuracy depends on the model used. In most livestock improvement schemes, selection is based on EBVs that are estimated using BLUP. A method used in prediction models for calculating the accuracy of selection on EBVs was presented by Wray and Hill (1989), based on selection index theory. A multiple trait extension of this method was presented by Villanueva et al. (1993), and Bijma et al. (2001) presented an extension for overlapping generations. These methods account for the reduction of genetic variance due to selection.

The optimisation of generation interval is automatically achieved when animals are selected on BLUP EBVs irrespective of their age and, indeed, the proportion of animals in each age class is itself optimised as a consequence of using BLUP selection (James, 1987). Animals of younger age classes have on average better EBVs, but also generally less variation in EBVs within age classes, the exception being when older age classes have been heavily selected,

and thus lower accuracy. The optimum proportion of selected younger animals depends on the difference in accuracy and on the genetic lag between age classes (i.e. on the genetic gain per year and the number of years) (Meuwissen, 1989).

Selection on BLUP EBVs maximizes response to selection. But response is only maximized with respect to the next generation. As BLUP uses information from the family members of an animal to predict its EBV, the members of that family have more chance to be jointly selected. This has more impact for lower heritability because more weight on the family information is given compared with the weight given to own performance, which leads in turn to more inbreeding (Quinton and Smith, 1995). Thus, long-term selection response is not necessarily optimised with BLUP selection of this effect on inbreeding compared to, e.g., mass selection (selection based on own phenotype). More inbreeding leads in the longer term to loss of genetic variation, and possibly inbreeding depression, and therefore reduces and offsets genetic gain. However, BLUP is usually superior to individual selection after 10 years of selection. Selection in the longer term can be optimised by its combination with some restriction on the average relatedness of the selected animals (Goddard and Wray, 1994; Meuwissen, 1997).

- **Debate on selection indices**

Selection indices should to be used with due attention to:

- a) Whether linearity of traits prevails. Where this is not the case, profit may be suboptimal. The profit function of some traits (e.g. % intramuscular fat, back fat thickness) is nonlinear.
- b) Breeders often have doubts whether the applied profit functions imply a profit for them or the retailers or the whole sector (national beef industry).
- c) The economic weights of the chosen traits and the corresponding indices could be invalid for some breeders as they could have different associated costs

due to e.g. the distance to the principal target beef markets which affects costs of transportation and inputs, production environment, quality of the soil which affects the quality and costs of forages, water and animal health, etc.

d) Selection indices could be invalid at a national level, besides the mentioned reasons on the previous paragraphs, due to the female genetic merit of herds that causes the breeder needs to include additional traits in the indices for the same breeding objective due to particular circumstances.

e) The different units in which profit could be expressed are not always suitable for all situations. For example, profit per unit of product minimizes the use of scarce resources for animal product and does not consider profit from increased output as genetic improvement if it is due to a change in the operation scale; this would not be the case if profit is expressed per unit of resource, such as per hectare.

For this last topic, see: van der Werf, 2007; Goddard, 1998; Shultz, 1990; Hovenier et al., 1993.

### **2.1.3. Evaluation of breeding program design**

In order to optimise genetic gain these factors must be taken into account and optimally related:

- selection intensity,
- selection accuracy,
- generation intervals, and
- genetic variance.

These factors are accounted in the classic Rendel and Robertson (1950) formula for predicting genetic response in the next generation:

$$R = \frac{i_m r_m + i_f r_f}{L_m + L_f} \sigma_g$$

where  $i$  is selection intensity,  $r$  is selection accuracy and  $L$  is generation interval. The subscripts  $m$  and  $f$  refer to males and females and  $\sigma_g$  is the genetic standard deviation.

Prediction of response is a key component in the evaluation of breeding program design. The methods used to evaluate breeding program designs can be classified into two groups: stochastic simulation and deterministic methods. Through these methodologies the expected selection response and inbreeding of breeding schemes can be predicted or estimated.

Deterministic methods are characterized to provide more insight and to be computationally fast, permitting a wider range of scenarios to be tested. On the other hand, stochastic simulation is more realistic and useful to validate deterministic methods, in particular because stochastic methods generally make fewer assumptions than deterministic methods, and these assumptions are often violated.

### **2.1.3.1. Deterministic and stochastic analysis**

Alternative breeding programs can be technically evaluated comparing predicted rates of gain and inbreeding. This can be done stochastically (e.g. Wray and Hill, 1989; Meyer and Smith, 1990; Brisbane and Gibson, 1995; Meuwissen, 1997; Meuwissen and Sonesson, 1998; Lewis and Simm, 2000; Grundy et al., 1998, 2000) or deterministically (e.g. Meuwissen, 1989; Goddard and Smith, 1990; Meuwissen and Woolliams, 1994; Meuwissen and Goddard, 1997; Meuwissen, 1997; Meuwissen, 1998; Bijma and Woolliams, 2000; Villanueva et al., 2000; Avendaño et al., 2003). It should be noted that while stochastic simulation is usually the easier to implement for realistic breeding

programs, deterministic simulation generally gives more insight, i.e. they are more illuminating in detecting the role of different components in genetic change.

With stochastic simulation, the breeding program is simulated in detail on a computer. Simulated records are generated for each animal by randomly sampling breeding values and phenotypes. This procedure can be performed for multiple traits and with selection in multiple stages for a set number of years and replicates, the number of replicates being dependent on the precision required. The distribution of phenotypes and EBVs varies according to the characteristics of the traits being examined.

After simulating the breeding scheme, the rate of gain and inbreeding can be quantified by analysing the simulated data.

Stochastic simulation can be very comprehensive due to fact that the breeding program is simulated in detail. At the same time, it has some advantages over a deterministic model. Many assumptions that have to be accounted for or ignored under deterministic models are implicitly handled by stochastic simulation, including reduction in variance due to selection, effect of multiple stage selection on variance, reduction in between family variance due to co-selection of sibs, reduction in selection intensity due to finite population size and the generation of non-normal distributions due to history of selection. However, the use of stochastic simulation has some disadvantages to evaluate breeding schemes. Stochastic simulation is time consuming, particularly when large populations are simulated. So it is less suited as an operational tool to quickly evaluate a number of alternatives. In addition, as stochastic simulation does not explicitly model factors such as accuracy, generation interval, etc, it can be difficult to extend the result to other breeding schemes that were not themselves simulated.

Stochastic simulation has also other uses like verification of deterministic models and testing of subroutines in animal breeding tools.

Deterministic methods, instead of simulating the breeding program on the individual animal level, use deterministic equations to predict gain and inbreeding (Woolliams J. et al, 1999, 2000).

In order to use deterministic methods it is necessary to derive the mechanisms determining genetic variance, gain and inbreeding. So it requires more insight into quantitative population genetics than stochastic simulation.

Advantages of deterministic methods include the much lower computational demand, so that many alternatives can be compared within limited time, and that these methods give more insight into gain and inbreeding in breeding programs because the mechanisms are modelled explicitly. However deterministic methods might not be precise if the mechanisms determining gain and inbreeding in populations are not properly modelled. Due to this, stochastic simulation may be used to validate and improve deterministic models.

#### **2.1.4. Changes in variance due to selection**

The genetic variance in selection candidates is equal to:

$$\sigma_g^2 = 0.25 \sigma_{gs}^2 + 0.25 \sigma_{gd}^2 + \sigma_{gm}^2$$

where  $\sigma_{gs}^2$  is the genetic variance in the selected sires and  $\sigma_{gd}^2$  in the selected dams and  $\sigma_{gm}^2$  is the Mendelian sampling variance. In an unselected and non-inbred population,

$$\sigma_{gs}^2 = \sigma_{gd}^2 = \sigma_g^2 \text{ and } \sigma_{gm}^2 = 0.5 \sigma_g^2$$

which means that sires and dams contribute each 25% to the genetic variance in an individual and that 50% of the genetic variance is due to Mendelian sampling. However, due to the Bulmer effect genetic variance is reduced by

selection in proportion to the reduction of phenotypic variance of the selected parents relative to their entire generation.

The genetic variance in the selected parents is thus equal to (Bulmer, 1971):

$$\sigma_{g(t+1)}^2 = (1 - 0.5 r^2 k) \sigma_{g(t)}^2$$

where  $k$  is the variance reduction coefficient calculated as  $i \cdot (i - x)$  ( $i$  being selection intensity,  $x$  is truncation point),  $r$  is the accuracy of selection and  $\sigma_{g(t)}^2$  is the genetic variance at the present generation.

The Bulmer effect leads to a reduction in genetic gain because the genetic gain is a direct function of the genetic variance. The variance reduction in the population is often close to 25%, which leads to a reduction in the absolute gain of about 13%. More importantly, however, the Bulmer effect reduces the genetic variance in parents (between family variance) but not the Mendelian sampling variance (within family variance). As a consequence, full and half-sib information becomes less important whereas information that includes the Mendelian sampling component of the selection candidate, such as own performance and progeny information, becomes relatively more important. These effects are important when comparing schemes that use different sources of information (e.g. progeny vs. sib information) (van der Werf, 2005).

### **2.1.5. Control of inbreeding and genetic diversity in breeding programs**

Genetic diversity refers to the amount of genetic variability present among animals in the population. It is influenced by the number of ancestors, the intensity of selection over time, bottlenecks in the population, genetic drift (random loss of genetic variation due to chance effects in segregation) and inbreeding (mating of related animals). Genetic variability is important because

of its impact in future response of selection, inbreeding depression (particularly in fertility and fitness traits) and the incidence of recessive genetic defects.

As mentioned, selection on BLUP EBV maximizes response to selection. However, response is only maximized with respect to the next generation. Long-term selection response is not necessarily optimised with BLUP selection.

It was also previously indicated that BLUP selection leads to more inbreeding than for example mass selection (selection based on own phenotype) due to the fact that BLUP uses information on all available relatives to estimate an animal's EBV. Using information from family members implies that members of the same (good) family have more chance to be jointly selected. With lower heritability there is relatively more weight on the family information (compared with the own performance), and the lower the heritability, the more BLUP selection resembles family selection and the more it leads to inbreeding. This is considerably higher with lower heritability. Loss of variance due to inbreeding is increased by increased use of relatives' information.

More inbreeding leads in the longer term to loss of genetic variation, and possibly inbreeding depression, and therefore reduces and offsets genetic gain.

This phenomenon has been well recognized and the accumulation of inbreeding is accelerated further with non-optimal use of advanced reproductive technologies such as MOET that allow for capitalization of elite genetics.

Increased genetic gain in the short term is usually associated with increased inbreeding which leads to decreased genetic gain in the long term, due to declines in fitness and genetic variance. Acceptable levels of inbreeding are difficult to determine, but different approaches point towards values around 0.5% and 1% per generation (Bijma, 2000).

Several studies have been carried out in order to show how short and long-term genetic responses are affected by selection on pedigree-based BLUP EBV (e.g. Toro et al., 1990; Leitch et al., 1994; Bijma et al., 2000; Henryon et al., 2009). In this sense, Bijma et al. (2001) developed a deterministic method to predict rates of genetic gain and inbreeding for livestock improvement schemes

with overlapping generations, BLUP selection, and progeny testing of male selection candidates, allowing balance short and long-term response. Also Daetwyler et al. (2007) discussed possible changes in rates of inbreeding and genetic gain when applying genome-wide selection. The accuracy of breeding values predicted on the basis of genomic information will be higher than for BLUP EBV. This increase of accuracy is due to better prediction of Mendelian sampling effects. The higher accuracy of Mendelian sampling effect estimates leads to lower co-selection of sibs, and therefore, reduces rates of inbreeding per generation. On the other hand, Clark et al. (2013) concluded that the use of genomic estimated breeding values (GEBV) increases genetic gain in optimal contribution selection. GEBV are more accurate and show more within-family variation than parent average EBV, which leads to higher genetic gains for the same restriction on inbreeding. Using genomic rather than pedigree-based relationships to restrict inbreeding provides no additional gain, except in the case of very large full-sib families.

So a key reason to maintain genetic diversity is to facilitate longer-term genetic gains, and this means that most breeding programs need to consider genetic diversity as well as shorter-term genetic gains. This is especially true for closed breeding programs for which future importation of competitive outside stock is unlikely.

To optimise selection in the longer term, it is useful to combine selection on BLUP EBV with some restriction on the average relatedness of the selected animals (Wray and Goddard, 1994; Meuwissen, 1997). This has increased the need for tools to monitor and control rates of inbreeding in a population, such as optimal genetic contribution selection (Meuwissen, 1997; Grundy et al., 2000). Methods and software tools have been developed that aim at describing inbreeding in a population level and to predict genetic contributions to maximize response to selection given a penalty on rate of inbreeding (Meuwissen, 1997; Meuwissen et al., 1998; Grundy et al., 2000; Kinghorn et al., 1997, 2000, 2002; Hinrichs et al., 2006).

In these approaches genetic gain is represented as predicted progeny merit:

$$x'G$$

and inbreeding can be constrained through parent coancestry. This is calculated as:

$$x'Ax/2$$

where  $x$  corresponds to the vector of selected parent or candidates contributions,  $G$  is the vector of genetic merit of selected parents and  $A$  is the numerator relationship matrix among selected parents, so the objective is to optimally balance the two elements, as will be addressed in this thesis.

The most generally useful strategy to maintain genetic diversity is to maintain variation across the genome, using methods that consider one or more of: population size, population structure, animal selections, mate allocations and information from genetic markers (van der Werf, 2009).

### **2.1.6. Implementing breeding program design**

Breeding program design can be implemented carrying out two different approaches:

- a) Strategic or rule-based approach
- b) 'Desired Outcomes' approach

#### **2.1.6.1. Strategic or rule-based approach**

In order to do a strategic optimisation it is necessary to set a whole population model of the breeding program in a deterministic manner. This model accommodates parameters from which these strategies emerge, e.g. which traits should be measured and in which animals, how many animals should be

measured and at what age this should be done, etc. Then, often, sets of rules are derived which indicate e.g. the optimal proportion of animals to be selected, the optimal age structure, the optimal system of mating, the optimal inbreeding rate, etc.

As examples of such rules can be mentioned:

- Do not mate males to females closer in relationship than cousin
- Assortatively mate selected males and females
- Cull females at 10 years of age
- Include performance or progeny testing to increase the accuracy of EBV estimates
- Select breeding animals across all age groups on BLUP EBV
- Use the 10% best male candidates mated to the 80% best female candidates.

Such rules are derived from generalised theories and concepts and these are usually not well integrated with each other. Thus, for example, theories and rules about selection, crossbreeding and inbreeding have been developed largely in isolation from each other, such that when they are mixed in real applications it is likely to miss the best overall strategy (Kinghorn and Shepherd, 1999). Also, this kind of rules is usually derived by optimising deterministic predictions of rates of genetic gain and inbreeding (Kinghorn et al. 2000).

After considering the most suitable design strategy for the breeding program the practitioners have to integrate the outputs of these rules in order to apply them. Each of these sets of rules is known to be appropriate on its own, but the best way to combine them in a real breeding program is using an integrating framework, typically Mate Selection, i.e. handling them simultaneously.

So, the rules-based approach results in a pre-defined design which dictates actions. In this way, the design of those breeding programs is fixed and breeding

alternatives can be compared by varying different parameters (van der Werf, 2005).

#### **2.1.6.2. 'Desired Outcomes' approach**

On the other hand, the Desired Outcomes approach is to let design emerge as the consequence of actions taken tactically, sometimes breaking what might otherwise be set up as inflexible design rules. Tactical optimisation is a more dynamic approach in which decisions are taken according to prevailing circumstances or opportunities (e.g. availability of feeding resources, current costs of reproductive inputs, sudden economic or financial situations, current or projected market prices, etc.) which can help give optimal selection results (Kinghorn et al. 2000), (van der Werf, 2005). It is necessary to remark that tactical optimisation operates at the level of individual animals, including decisions on which animals to select, collect semen or oocytes from and which males to mate to which females. However, a tactical approach does not provide information on how trait recording and resources could be changed to improve genetic gain in the future.

The tactical approach to breeding is driven by specifying desired outcomes. Any one solution that is predicted to best suit the breeder's needs is a set of actions for which the outcomes can be predicted (e.g. genetic merit of progeny, cost of semen purchase, number of mating paddocks required, etc.) (Kinghorn, 2005).

Some tools like BLUP (see previous section 2.1.2.3) or Mate Selection (see later section 2.4) can be used in order to take tactical decisions for selection.

In order to achieve these results special techniques or algorithms suitable for global numerical optimisation are proposed, among which can be highlighted those related to Evolutionary Computation (see later section 2.3.1).

## ***2.2. Breeding program design for beef cattle production systems***

The comparison of the economic values used in producing an Index across studies is difficult due to the diversity of production systems. These differences in production systems may happen within the same country. In the same way, assessing objectives and developing standard indices becomes a difficult task when there are differences in climates, biotypes, feedstuffs and markets in the USA (Harris D., 1998). Similar conclusions were made by Phocas et al. (1998) who showed the existence of substantial variation in economic values between different beef production systems in France. Moreover, when comparisons of breeding objectives are made between countries, difficulties increase according to the even greater diversity of production systems.

Australian authors have focused their studies on the development of breeding objectives for purebred beef cattle (e.g. Graser et al., 1994). On the other hand, studies of authors from Europe are oriented to dual purpose (milk and meat) production systems (e.g. Phocas et al., 1998). On the contrary, in the USA crossbreeding with feedlot finishing production systems are more widespread than in Australia.

Many studies about breeding objectives for Australian conditions have been based on BREED-OBJECT (Schneeberger et al., 1992; Barwick, 1993), a computer program designed to provide customized indices for beef producers, and have focused upon increasing the accuracy of selection for the overall breeding objective.

In Argentina breeding objectives for purebred beef cattle are mainly oriented to beef production for the domestic market on grass-fed production systems.

### **2.2.1 The use of Young Sire Programs**

A Young Sire Program (YSP) is designed to identify suitable elite young sires, encourage their use and prove them quickly to achieve widespread usage. This will accelerate genetic gain and provide elite sires for both domestic and international marketing.

A YSP involves mating young sires to generate test progeny. One of the main objectives of a progeny test is to accurately test the young sires to use in a breeding program. Young sires compete with proven sires so only the best young sires are selected to be mated in the nucleus herd. Using young males reduces the generation interval, which in turn contributes to the acceleration of the rate of genetic gain. Progeny testing could be an expensive exercise for individual studs. Thus, joining a cooperative YSP appears as a suitable solution for stud-farms with a limited herd size, decreasing the cost of genetic progress compared with the use of imported semen, and commercial clients benefit from more accurate breeding values and better genetics (Banks et al., 1998).

Several YSP are been carried out by different companies and breeder associations in Australia, U.S.A. and other countries, not only for beef cattle but also for dairy cattle and other species. All of them are based on progeny test in order to test and select the best young sires. Among these YSP could be mentioned Meat Elite Australia (MEA) for meat sheep (which is carried out by an Australian sheep breeding group) and the one carried out by Angus Australia which started in 2004. MEA reports that it is achieving 21% faster genetic progress than the average for all terminal sire groups across Australia for a range of profit-driving traits, although this benefit probably involves other aspects of the breeding program, such as accurate recording. On the other hand, Angus Australia has the objective of doubling the rate of genetic gain within the following 10 years, i.e. before or in 2014. Apart from these two groups, the others do not report their rate of genetic gain estimates.

In this thesis different aspects of the implementation of a YSP are analysed through a deterministic modelling and simulation (in Chapters 5) that was also based in progeny testing, and by stochastic simulation (in Chapter 6). In Chapter 5, a genetic response equal to 0.200 TBV genetic standard deviations per year of mating was obtained in the traditional approach (no-LAMS) and around 3.0% more genetic gain was obtained applying a look-ahead mate selection scheme approach (LAMS) compared to no-LAMS when 100 proven bulls are available as candidates to be selected as sires. While in Chapter 6, the average increase in the TBV mean per year of mating was 0.289 for LAMS and 0.270 for no-LAMS, i.e. 7.04% higher for LAMS than for no-LAMS. Also it appears that after the 14th year of mating the achieved TBV mean is slightly higher in LAMS than in no-LAMS ( $\approx 0.303\%$  at the 20th year of mating), but without statistical significance. Even though the differences between LAMS and no-LAMS are small, they appear higher in the analysis done in Chapter 5 because sampling error is not accounted for.

A fair comparison between the results in these Chapters of this thesis and the estimations of the rate of genetic gain in the two mentioned YSP carried out in Australia would not be feasible as in MEA the reported percentage genetic gain comes from the difference between a subpopulation and the rest of that population and not in a time basis and, in the case of Angus Australia, the stated objective does not appear to be based on a prediction of the impact of the YSP per se.

### ***2.3. Genetic algorithms and their application in animal breeding programs***

Genetic algorithms have demonstrated much utility when these are applied to breeding program design (as shown by Kinghorn, 1998; Meszaros et al., 1999; Kinghorn et al., 2002).

As it can be observed from the previous topics, Mate Selection becomes a complex task, even more as the number of variables involved in the breeding program increases (number and size of herds, number of traits considered, type of costs, etc.). In a rules-based approach we *derive* a solution to be followed, however in a tactical optimisation such as Mate Selection, we have to *find* a good solution. To solve this, special techniques or algorithms suitable for global numerical optimisation are used. One of the most efficient of these algorithms is Differential Evolution.

### **2.3.1. Evolutionary Computation**

This field belongs to a more extended area of computer science named artificial intelligence. It is aimed at solving complex optimisation problems using iterative processes to search for the optimal solution in order to achieve a desired result. Although the search of the solution space follows a systematic process, randomness is exploited in a similar manner to biological evolution due to the fact that Evolutionary Computation uses stochastic methods.

Optimisation consists of choosing the solution from a set of available or feasible alternatives that maximizes or minimizes a specific objective function (i.e. the optimal solution). Through the objective function each candidate solution is measured and its fitness is obtained. Based on this fitness a candidate solution is chosen or rejected according to how well it fits the proposed model of the problem and the method applied for the evaluation of this fitness. Sometimes the objective function is nonlinear and contains many real variables. In these cases it determines a tortured or multidimensional solution space, i.e. with many peaks and valleys, making it much more complicated to reach a global optimal solution, which is to reach the highest point in the highest peak or the lowest point in the deepest valley depending on if it is desired to maximize or minimize the objective function (see Figure 2.1). It is hereafter assumed that functions are to be maximized for simpler presentation. Most Evolutionary Computation

algorithms are suitable for global numerical optimisation (Storn and Price, 1997).

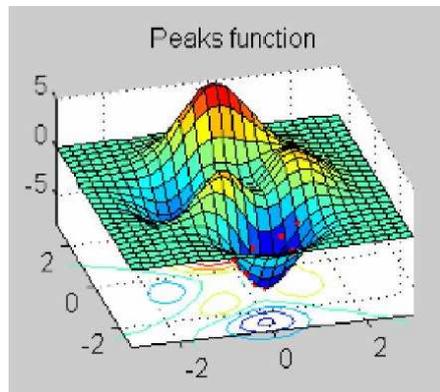


Figure 2.1. Multidimensional function space. Global optimisation  
From R. Storn and K. Price, Differential Evolution Homepage

The name “evolutionary” is due to the fact that the techniques applied by these algorithms emulate the biological processes of evolution. In this way, they use mechanisms analogous to selection, mutation and recombination (crossing over) in repeated cycles, not being only constrained to replication and propagation. These mechanisms are called search operators and their aim is to provide variability modifying the candidate solutions in order to explore the solution space.

Evolutionary Computation comprises techniques or methods that use heuristic algorithms, including Evolutionary Programming, Evolutionary Strategies, Genetic Programming, Genetic Algorithms and Learning Classifier Systems. Other algorithms like Differential Evolution and Artificial Life are loosely related to Evolutionary Computation.

"Heuristics" comes from the Greek εὑρίσκειν, *heuriskein*, which means "to find", i.e. these are methods to solve computational problems through search procedures, coming to a solution which is at or close to the best possible answer (optimal solution).

These techniques have applications in very varied areas like automated design, control engineering, economics, robotics, software engineering, bioinformatics, etc.

### **2.3.2. Differential Evolution**

This method was developed by Rainer Storn and Kenneth Price and published in 1997. This is aimed at carrying out numerical global optimisation, being significantly fast, robust and simple. One of its main features is that it uses a floating-point representation which allows a more precise solution to be reached. It can be described as a population-based stochastic function maximizer (Storn and Price, 1997).

The Differential Evolution (DE) algorithm works using 3 factors which control evolution:

- $NP$  (Number of population vectors) which is the population size,
- $F$  (select weighting factor) which manages mutation and
- $CR$  (crossover constant that mediates the operation).

DE works with populations formed by parameter vectors that describe solutions, which are real-valued through the applied objective function. It should be mentioned that these vectors can also be coded as integers wherever this is more suitable for the application involved.

The DE algorithm (see Figure 2.2) involves challenging each one of the vectors of the current population (target vectors) against another vector (the trial vector) generated by evolution processes. The vector which overcomes this challenge will constitute a new vector of the new population for the next generation.

Schematically its steps in each generation are:

- First a target vector is chosen.
- On the other hand, the weighted (by the  $F$  factor) difference between two other vectors (vectors  $a$  &  $b$ ) is added to a third vector (vector  $c$ ), all these vectors being randomly chosen from the same current population. The resulting vector (the mutant) becomes a part of a second population (mutant population). So the mutant vector is achieved computing  $c + F * (a - b)$  for each parameter. Generally  $F$  is set between 0.2 and 2.0.
- Then a trial vector is generated by a nonuniform crossover operation between the target vector and the mutant vector.
- Finally, if the value of the trial vector is better than or equal to the target vector it is joined to the new population for the next generation. Otherwise, the target vector retains its position into the next generation.

It should be noted that the parameters of the trial vector are crossed over coming from the target and the mutant vectors in different proportions according to the value of the  $CR$  constant probability. Thus,  $CR$  ranges between 0 and 1.

$CR$  and  $F$  settings can be modified along the execution of the DE session in order to explore the solution space in a suitable way. Higher values of  $F$  permit wider searching of the solution space and are more convenient during the first iterations. Then it is better to use lower values to be more conservative and to speed convergence, changing periodically for some iterations (generations) to not get stuck in a relatively small place (Mayer et al., 2005). On the other hand, as for  $F$  values, it is convenient to start with higher values of  $CR$  and then to set it to intermediate or even lower values. But, in general, it should be said that to set these factors properly needs some investigation for the type of problem to be solved.

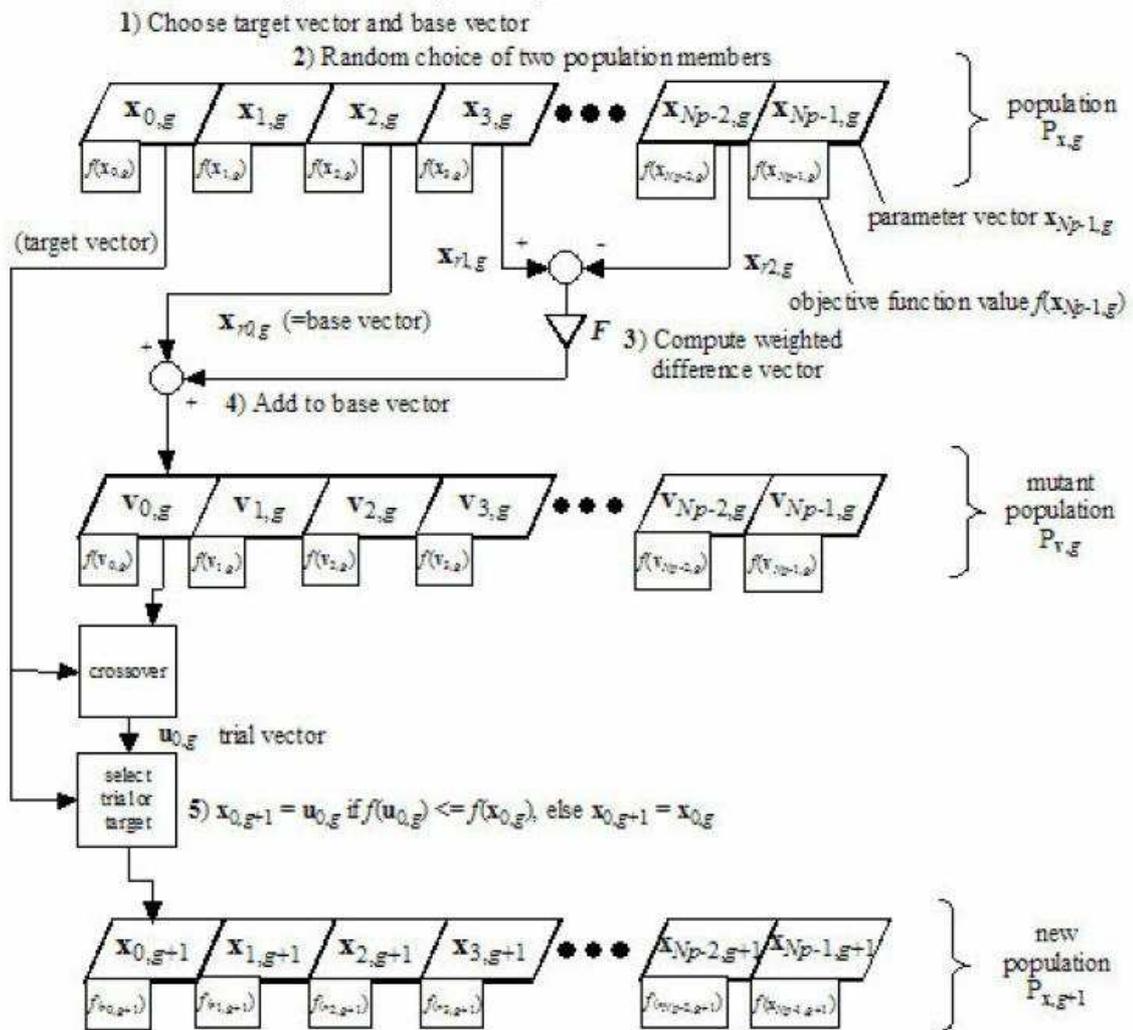


Figure 2.2. Differential Evolution scheme

From R. Storn and K. Price, Differential Evolution Homepage

### • Mutation in DE

The aim of mutation is to inject or maintain diversity in a population.

The strategy of computing vector differentials (vector  $a$  - vector  $b$ ) affected by the weighting factor ( $F$ ) in the DE algorithm assures the efficient exploration of the solution space (Storn and Price, 1997), with 'automatic' adjustment of mutation levels separately for each parameter according its level of consistency among the prevailing solutions, as noted above. This differential

vector is added to another randomly chosen vector of the current population vector in order to obtain a new mutant vector.

- **Recombination in DE**

Recombination (also called crossover) is an operator by which the parameters of different vectors (two or more) are combined into a new vector. In this way, the most successful combinations are kept so the search is directed to exploring the most promising areas of the solution space (Storn and Price, 1997). When recombination is used with mutation, algorithms become faster and more robust.

As mentioned, in the DE algorithm a target vector is randomly chosen every generation from the current population in order to be recombined with the mutant vector. In this way a trial vector is obtained. The parameters that constitute the trial vector will come from the target or the mutant vector according to the value of the *CR* crossover constant and a randomly uniformly distributed number. If the value of the latter is greater than *CR* the trial vector gets its parameter from the target vector; if not it gets the parameter from the mutant vector (Storn and Price, 1997). Unlike other genetic algorithms, DE uses nonuniform recombination to build the trial vector, i.e. that parameters to build it could come from the target or the mutant vector in different proportions.

- **Selection in DE**

In this algorithm a random approach is used for the selection mechanism. This is done by the competition between each trial vector and one of its parents, i.e. the target vector. If the former is better than or equal to the latter, it will constitute the corresponding population member in the new generation (Storn and Price, 1997).

After the last target vector is challenged the current population is swapped with the new population in their roles and a new cycle starts. These cycles continue until convergence has apparently been reached (Storn and Price, 1997).

- **Applications**

DE has nowadays many applications. Some of them are: Multiprocessor synthesis, Neural network learning, Digital filters design, Crystallographic characterization, Synthesis of modulators, Scenario-Integrated Optimisation of Dynamic Systems, Differential Evolution for Multi-Objective Optimisation, Radio Network Design, Reflectivity Curve Simulation, etc. (Storn and Price, Differential Evolution, <http://www.icsi.berkeley.edu/~storn/code.html>). In this thesis, DE will be applied to Mate Selection.

#### ***2.4. Mate Selection***

After establishing the breeding objectives, individual mating allocations need to be made to efficiently meet them. A Mate Selection program establishes which animals to select as parents (i.e. animal selection) and which sire should be mated to which dam (i.e. mate allocation) in order to achieve that goal.

In an effort to predict breeding values as accurately as possible, methods that utilise family information have found widespread use. Selecting on selection index or animal model predictors of breeding values results in increased coselection of sibs, and thus increased rates of inbreeding (Falconer and Mackay 1996), as noted previously. This consequence takes more importance with traits of low heritability as there is relatively more weight on family information than own performance information (Quinton and Smith, 1995).

When breeding schemes within populations are designed, increasing inbreeding levels need to be avoided if maximization of genetic gains is desired in a longer term period. In order to achieve this, these directions have to be indicated in the objective function in some way.

Some authors have designed selection methods that reduce rates of inbreeding. Grundy and Hill (1993) and Verrier et al. (1993) reduced the weight of the family mean in their selection index relative to that in BLUP EBV, which reduced the probability of coselection of relatives and thus reduced inbreeding. Brisbane and Gibson (1994) and Wray and Goddard (1994) selected animals while putting a cost on the average relationship of the selected animals. Meuwissen (1997) and Meuwissen and Sonesson (1998) applied a dynamic iterative method to limit the future contributions for inbreeding of the selected animals. Later, Hinrichs et al. (2006) have improved the method proposed by Meuwissen and Sonesson reducing the computational time by the use of an alternative algorithm to calculate inbreeding relationships among selected parents for large databases. In addition, Pong-Wong et al. (2006) developed an approach that guarantees the finding of optimum genetic contributions of candidates to control inbreeding in the offspring generation using semidefinite programming, as the dynamic iterative methods mentioned above can sometimes find incorrect results. These methods mentioned above calculate optimal genetic contributions. However, it is also possible to search for optimal contributions, typically using an evolutionary algorithm.

An advantage of using an EA in this context is that it can be simply extended to accommodate objectives other than optimal contributions for gain and diversity. Decisions for selection of parents and the corresponding mating allocations can be affected by many other factors. Among them can be considered inbreeding on progeny, the availability and costs of reproductive mating technologies (natural mating, MOET, AI, IVF, etc.) and, furthermore, factors like targeting of gene marker frequencies, connections between herds, multi-stage selection, stock migration, reproduction costs, levels of investment,

etc. (Kinghorn et al., 2009). The Mate Selection strategy could also lead to divergent lines serving different segments in the client market.

In this way, Mate Selection becomes a complex task, even more if the number of variables involved in the breeding program increases (e.g. number and size of herds, maximum number of matings, number of traits considered, type of costs, number of considered logical constraints, etc.). ‘Strategic or rule-based’ or ‘Desired Outcomes’ approaches (see previous section 2.1.6) can be carried out to solve this. The choice of which traits to measure or to genotype on which animals is the only key issue not covered by the combined Mate Selection approach. This has to be handled by extension to the method.

As mentioned, to do a strategic optimisation it is necessary to set a whole population model of the breeding program, which accommodates parameters from which emerge these strategies. While tactical optimisation is a more dynamic approach in which decisions are taken according to prevailing circumstances or opportunities which can help give optimal selection results (Kinghorn et al. 2000), (van der Werf, 2005). Mate Selection can be used in order to make tactical ‘implementing’ decisions, by specifying desired outcomes.

In order to achieve these results in the tactical approach special techniques or algorithms suitable for global numerical optimisation are proposed, among which can be highlighted those related to Evolutionary Computation.

As previously mentioned, Evolutionary Computation comprises algorithms that emulate the biological processes of evolution. This means that they use mechanisms analogous to selection, mutation and recombination (or crossing over) in repeated cycles or generations. Through these evolution processes the algorithm makes populations evolve creating candidate solutions which are selected or rejected according to an objective function (Bäck 2003). One of the most efficient of these algorithms is Differential Evolution.

It is necessary to consider how to present the problem to the applied algorithm. An appropriate choice of representation for the populations is crucial

for EC and largely depends on the nature of the problem (Gondro and Kinghorn, 2007). Applying the DE algorithm, Kinghorn and Shepherd (1999) presented a procedure by which a vector of numbers can be transformed to derive a pattern of animal selection and mate allocation. Thus, the optimum solution is achieved when the vector which maximizes the objective function is found.

#### **2.4.1. Objective function: the Mate Selection Index**

In this way Mate Selection is driven by the desired outcomes specified in an overall objective function. This objective functions accommodates the different technical and logistic components that are desired. There can be very many components, including, just for example, progeny merit and funds spent on semen. From the component outcomes evaluated for each mating set the corresponding Mate Selection Index (MSI) is calculated, this being value under the prevailing objective function. In order to be evaluated, these components should ideally be expressed on the same scale, which is generally the scale of the multitrait breeding objective. It should be noted that this last requirement is practically never achieved as it is too difficult to calculate how to weigh the different components. Hence, the importance of the desired outcomes approach.

Thus, the MSI is a function of the resulting mating set composed of the total selected parents among all the mating candidates and the specific mate allocations between them (Kinghorn et al., 2009).

As previously mentioned, the objective function should take into account all the considered factors and should be able to be changed if a different outcome is desired. The need for this change is brought about by the fact that the result of using any one objective function cannot be predicted, and must therefore be tested in the system. The practitioner can then change the objective function (and thus MSI) in a manner considered to be required to satisfy desires. This may require several iterations. This process gives the opportunity to the

breeder not only to achieve an optimal solution but also a solution that “satisfies” her/his desires, with some ‘ownership’ of the solution.

In order to accommodate antagonistic objectives it is necessary to make trade-off decisions, simultaneously balancing the science applied and the experience and preferences of the breeder. In this way “satisfaction” becomes a very important aspect in the Mate Selection process and the final achieved solution will be surely implemented and will not only be taken just as a recommendation (Kinghorn et al., 2002; Gondro and Kinghorn, 2007; Kinghorn et al., 2009).

The final outcome of the overall Mate Selection process is a mating list which condenses almost the total key issues in animal breeding (i.e. operational and management, genetic, health and financial) into a tactical set of actions for the breeder, simultaneously targeting all components in the overall objective (Kinghorn, 2000; Kinghorn et al., 2002).

Some systems are currently available for breeders of different species to design their own breeding programs using Mate Selection. Between them it can be mentioned TGRM (Total General Resource Management, (X’Prime, [www.xprime.com.au](http://www.xprime.com.au)), MateSel in Pedigree Viewer (Kinghorn, [www-personal.une.edu.au/~bkinghor/pedigree.htm](http://www-personal.une.edu.au/~bkinghor/pedigree.htm)), Ani-Mate (Amer, [www.abacusbio.co.nz/products.html](http://www.abacusbio.co.nz/products.html)), GENCONT/OCSELECT (Meuwissen, 2002; Hinrichs et al., 2006) and EVA (Berg et al., 2006).

## **CHAPTER 3**

# **OPTIMISING KEY OUTCOMES USING MATE SELECTION**

### ***3.1 Summary***

Optimal balance between high genetic gain and low parental coancestry when selecting breeding animals can be achieved with the optimal set of contributions of the selected individuals' genes to the breeding program. These optimal contributions can be found by using a deterministic approach (Meuwissen, 1997) or a stochastic search algorithm, such as a Differential Evolution algorithm. The latter approach accommodates the simultaneous management of factors other than high genetic gain and low parental coancestry. Moreover, attention to mate allocation, using simultaneous Mate Selection, adds potential to manage many additional factors.

The objective of this chapter was to illustrate the importance and interactions of the various factors affecting mate selection. This has been achieved by determining the optimal set of matings from a list of candidate parents under different policies, constraints and parameter settings using an example dataset of a seed stock herd of beef cattle with overlapping generations. Among the considered policies considered were the maximization of progeny merit, the reduction of short and long-term inbreeding, targeting variance of merit in progeny and a preliminary general test of Look Ahead Mate Selection approach (LAMS), which, as it was already mentioned, targets grand progeny merit by maximizing the genetic selection differential in the progeny generation (following Kinghorn and Shepherd, 1994 and Hayes et al., 1998).

In all the analysed cases a reduction in the levels of coancestry corresponded to a reduction in genetic gain in the next generation. At the same time, these effects were associated with a reduction in the levels of short-term inbreeding, an increment of the variance of the progeny merit in the next generation and an increment in the number of sires selected as parents (and consequently lower relatedness among their progeny).

The application of a Mate Selection tool can lead to sustainable genetic progress, with maintenance of genetic diversity and avoidance of the undesirable consequences of inbreeding. Adding other technical and logistical factors allows practical implementation of sustainable breeding programs.

LAMS approach has resulted in some compromise in slightly lower parental coancestry and lower progeny merit in the next generation, but caused an increment of the variance of the progeny merit in the next generation. This increment allowed bigger differential selections and, in consequence, bigger genetic gains in the second generation, i.e. the predicted grand progeny merit. The results of this general test show some promise for the LAMS approach, and it will be further tested for specific scenarios in Chapters 5 and 6.

### ***3.2 Introduction***

Animal breeding aims at the genetic improvement of livestock populations in order to produce the desired animal products in an efficient and sustainable way. To achieve this goal it is necessary to select the most suitable individuals in the current generation and mate them in such a way that a new generation is obtained in which is observed a genetic response in the desired direction. This is in fact Mate Selection – the actual implementation of a breeding program. What matters is finding the best Mate Selection outcomes, not through applying rules, but through evaluating the predicted consequences of Mate Selection solutions, and finding the best one.

When breeding schemes within populations are designed, increasing inbreeding levels need to be avoided if maximization of genetic gain is desired in a longer term period. Optimal balance between high genetic gain and low parental coancestry when selecting and mating breeding animals can be achieved with the optimal set of contributions of the selected individuals' genes to the breeding program. These optimal contributions can be deduced by

numerical solving (Meuwissen 1997). Other issues are also impacted by the pattern of mate allocations such that optimal mate selections are required. These optimal mate allocations can be found by a search strategy such as simulated annealing (Fernandez J., Toro M.A., Caballero A., 2001) or evolutionary algorithms (Kinghorn et al., 2002).

Decisions for selection of parents and the corresponding mating allocations can be affected by more factors besides the maximization of progeny merit, genetic diversity levels and restriction of progeny inbreeding and long-term inbreeding. Among them can be considered, for example, the search of technical aspects like the improvement of the connection between herds to achieve more accurate genetic evaluations, the avoidance of deleterious recessive genes or others like the present and future conditions of the beef market, and the desires of the breeder with respect to a specific trait in the breeding objective.

Thus, very different factors have to be taken into account in the design of a mating program, making Mate Selection a complex task, as described by Kinghorn et al. (2002). An optimisation model based on an evolutionary algorithm gives the flexibility to consider all these factors simultaneously by including them in the overall objective function. Although the importance of the various factors related to mate selection have been described in general terms, it may be not always clear how they interact and to what extent they are interchangeable. For example, it would be relevant to know whether managing long-term inbreeding would lead to a very different choice of managing short-term inbreeding. Similarly, when taking into account not only the mean but also the variance in future progeny, would that affect the short-term optimal response? The effects of the various population and genetic factors affecting mate selection and their interaction could be explored in an example data set reflecting a realistic situation in beef cattle.

The objective of this chapter was to illustrate the importance and interactions of the various factors affecting mate selection. This was achieved by

determining the optimal set of matings from a list of candidate parents under different policies, constraints and parameter settings using an example dataset of a seed stock herd of beef cattle. In particular, optimisation of progeny merit, coancestry, progeny inbreeding, variance in progeny and performance in later generations were considered.

### ***3.3 Materials and Methods***

#### **3.3.1 Proposed scenarios**

Different factors were considered, mainly related to the genetic features of the herd. I will first describe these factors and subsequently describe the simulation of a dataset and a systematic way to compare outcomes by varying emphasis on the various factors in the mate selection optimisation. The factors to be considered are:

- a. Maximization of progeny merit
- b. Minimization of long-term inbreeding (via parental coancestry)
- c. Minimization of short-term inbreeding
- d. Variance of merit in progeny
- e. Look Ahead Mate Selection, targeting grand progeny merit

Giving more or less weighting or emphasis to one of the parameters genetic gain, short-term inbreeding or coancestry caused an impact on the others, as is the case for altering weightings in a multi-trait selection index. A further description of the various factors follows:

- **a. Maximization of progeny merit**

A common objective of breeders is the maximization of the progeny additive genetic merit in the next generation. Then, selection of parents based on their EBVs is carried out in a way to achieve that aim. Under additivity, the actual mate allocation of selected parents does not affect the predicted progeny genetic merit, given that the number of matings per candidate has been set. These settings should be considered as prevailing conditions more than arbitrary rules. In this way, EBV selection is tactical, i.e. carried out without applying any arbitrary rule.

This example aims to show how different genetic parameters are affected in the population when the maximum possible genetic merit is achieved in the next generation.

- **b. Minimization of long-term inbreeding**

Long-term inbreeding is connected to the effective population size ( $N_e$ ) and genetic diversity (see 2.1.4 and 2.1.5 sections above).

Coancestry describes the strength of relationship between animals in a breeding population. The coefficient of coancestry between individuals A and B is the probability that two gametes taken at random, one from individual A and the other from individual B, carry alleles that are identical by descent, i.e., are identical and derived from the same common ancestor. The coancestry of two individuals is equal to the inbreeding coefficient of their progeny. This relatedness affects the genetic diversity, so more coancestry indicates less genetic diversity. At the same time, high coancestry leads to high long-term inbreeding.

The average of estimated parental coancestry depends on the number of animals selected as parents, the extent to which each will be used, the relationships among them, and hence the depth of pedigree included.

Animal model BLUP gives the most accurate estimates of EBVs, making use of information on all known relatives. Selection on EBVs tends to result in

the selection of related animals. Thus, the greatest short-term genetic improvement is achieved but at the same time the loss of gene diversity in later generations is accelerated, i.e. the effective population size is reduced and, therefore, the long-term breeding potential. This tendency is increased when selection is for a trait with low heritability where information on family members is relatively more important compared to information on the individual and consequently more members of the same family are selected.

The loss of diversity leads to less ability to select better animals, and as a consequence a decrease in genetic progress.

- **c. Minimization of short-term inbreeding**

Short-term inbreeding ( $F$ ) refers to the average coefficient of inbreeding achieved in the next generation (see 2.1.4 and 2.1.5 above).

Under complete dominance the effect of recessive genes can only express when they are in a homozygous state. The higher inbreeding level indicates the higher probability that recessive genes get this state and, therefore, can express. Often these recessive genes are lethal or cause serious problems. So the rise of the inbreeding level leads to what is called inbreeding depression: a decrease in production (e.g. slower growth), reproduction (e.g. lower conception rate and progeny survival) and overall fitness (e.g. animals are more likely to have health problems). A similar argument can be made for partial dominance.

In addition, inbreeding causes individuals in a population to be genetically more similar, so genetic diversity tends to be lower, leaving less room for selection.

When breeding schemes within populations are designed, increasing inbreeding levels need to be avoided if maximization of genetic gains is desired in a longer term period. Average inbreeding and its effects can be reduced in the progeny generation by avoiding the mating of relatives (Falconer and Mackay, 1996). However it should be noted that this policy is generally ineffective for

reducing long-term inbreeding as the latter depends mainly in the number of parents selected and their genetic relationships.

- **d. Variance of merit in progeny**

Assortative mating takes place when sexually reproducing organisms tend to mate with individuals that are like themselves in some respect (positive assortative mating) or dissimilar (negative assortative mating). Assortative mating leads to genetic covariance between mates which increases or decreases depending on whether assortative mating is positive or negative.

The assortative mating of preselected parents according to their EBVs can change the progeny genetic variance. It should be mentioned that actually in this case, where Mate Selection is applied, assortative mating is not invoked directly but induced indirectly by targeting increased variation in progeny. The change in the progeny genetic variance depends on the heritability of the trait, the genetic correlation created between mates and the selection intensity (Bulmer, 1971), (Falconer and Mackay, 1996). High selection intensities narrow the range of the trait values and therefore reduce the potential of positive assortative mating to increase the genetic covariance between mates.

Under positive assortative mating, greater long-term gain arises because greater genetic selection differentials in later generations become possible, which result from the relatively larger progeny genetic variances that are produced with positive assortative mating relative to random mating (Allaire, 1993).

- **e. Look-Ahead Mate Selection**

As was previously mentioned, in general breeders look for maximizing the progeny additive genetic merit in the next generation. For this, they select males and females as parents based on their EBVs and mate them accordingly.

Frequently, as it was previously analysed, breeders can consider additional aspects during the selection and mating process like reduction of inbreeding and coancestry or the advantages of assortative mating.

Kinghorn and Shepherd (1994) have proposed an alternative approach, i.e. Look-Ahead Mate Selection, in which the aim is to maximise the predicted genetic merit of the grand progeny of selected parents. This is achieved by maximizing the predicted genetic selection differential achieved by selection in the progeny generation. Note that the evaluation of individuals is not delayed by a generation, but a prediction of second generation impact is immediately available such there is no deleterious impact via longer generation intervals. This approach takes advantage of increased identifiable genetic variance in the progeny generation, which can arise from, for example, better genetic connections among fixed-effect groups in the progeny generation, or the use of positive assortative mating. Assortative mating effects are improved by suitable mate allocations which yield better information from the pedigree structure. Connection between different fixed effects groups leads to a better comparison of animals and more accurate genetic evaluations. This procedure may imply the selection as parents of animals with the only benefit to improve the information from the pedigree structure and the connectedness between different fixed effects groups. Thus, the average genetic merit of progeny may in fact be compromised in the first generation, compared to aiming only at that generation, but will be greater in the second generation, compared to running normal mate selection in tandem.

It should be mentioned that sensible treatment of parental coancestry across the population might impact appropriately for increasing genetic gains over a wide range of future generations. However, considering the topics that will be investigated later in this thesis, the actual specific interest relates to impact on gains in a specific future generation (i.e. generation 2) through a close technical prediction of explainable variance of parental mean EBV versus residual genetic variation in groups of progeny from known parents.

Thus, LAMS will be illustrated as a simple case to set the scene for investigating the setting up of a Young Sire Program using Look-Ahead Mate Selection later in Chapters 5 and 6.

### **3.3.2 Data simulation**

A single replicate of a dataset of a seed stock herd of beef cattle with overlapping generations was generated using PopSim (<http://www-personal.une.edu.au/~bkinghor/genup.htm>).

A base dataset of 10 sires and 250 dams was generated, where each sire was mated to 25 dams and each dam having one progeny. Foundation animals were assumed to be unrelated, unselected and randomly sampled from an infinite population. It should be noted that the base dataset has not undergone Mate Selection. Phenotypes of 12-month weight with mean equal to 400 kg, std dev 36 kg and heritability 0.3 were simulated. First progeny were born when parents were age 2 years, culling for age after 4 and 8 mating cycles for males and females respectively with random adult annual survival of 90%. Selection was on BLUP EBVs for 10 mating seasons. Several mating seasons are necessary to develop the final reproductive structure of the simulated population, allowing the replacement of sires and dams according the set ages for culling and the selection criteria and the development of the pedigree structure of the population (genetic relationship among animals). In this way, a population with overlapping generations is obtained and its final genetic parameters can be estimated.

After the last mating season the simulated population had a total of 2693 animals, of which 883 were available for reproduction (341 males and 542 females), each one with its own 12-month weight BLUP EBV. Based on their EBVs, the top 10% of males and 80% of females of these latter were considered as candidates for selection, resulting in totals of 34 males and 434 females with a requirement to make 250 matings. The maximal contribution was equal to 25

per male, while females had only 1 allowed mate as embryo transfer was not considered.

These settings should be considered as prevailing conditions more than arbitrary rules.

### **3.3.3 Analysis method**

Following this structure, the Matesel option in the Pedigree Viewer system (<http://www-personal.une.edu.au/~bkinghor/pedigree.htm>) was then applied to determine the set of matings (including optimum contribution of candidates) to generate the next set of progeny. The source code of the Matesel system was modified by the author in order to achieve the custom analyses required in this chapter.

The Matesel system can balance progeny index response (genetic gain) and parental coancestry in several ways, targeting a selected number of degrees, where zero degrees corresponds to the maximum genetic gain and 90 degrees corresponds to maximum genetic diversity (see Figure 3.1).

Matesel has a number of settings that can be used to manage the solution in a particular region of the solution space either, for example, above or below a selected number of degrees or constraining the solution to a maximum or a minimum value for specific genetic parameters such as progeny merit and coancestry. In this way, during the analysis process, Matesel penalises those solutions that do not comply with these settings or are out the desired solution space limits. Among these settings Balance Strategy, which relates to management of the balance between genetic gain (predicted index value in progeny) and parental coancestry, and different Weightings (e.g. on coancestry, on progeny inbreeding, etc.) are found.

In all the analysed cases Balance Strategy was set to 3, which put a hard constraint on the selected number of target degrees (in red figures, see Figure 3.1), severely penalising solutions that break the target degrees constraint. For

target degrees less than 45 Matesel aims to maximise genetic gain in progeny, maintaining the solution at or above the given value of target degrees, while, for target degrees greater than 45, aims to minimise coancestry, maintaining the solution at or below the given value of target degrees (see Figure 3.1).

Given the constraints on sire uses and the number of females being mated, a frontier of genetic gain and coancestry can be established. For any one chosen Balance Strategy 3 policy on balance between genetic gain and coancestry, the corresponding point on this frontier line indicates the best outcome in terms of high genetic gain and low coancestry, with each such policy characterised by a number of degrees. Along each of the points of this frontier an optimal mating list is set according the corresponding levels of genetic gain and coancestry. During the analysis process a trailing path shows the past progress of the differential evolution algorithm towards the selected point (Newman et al., 2009) (see Figure 3.1).

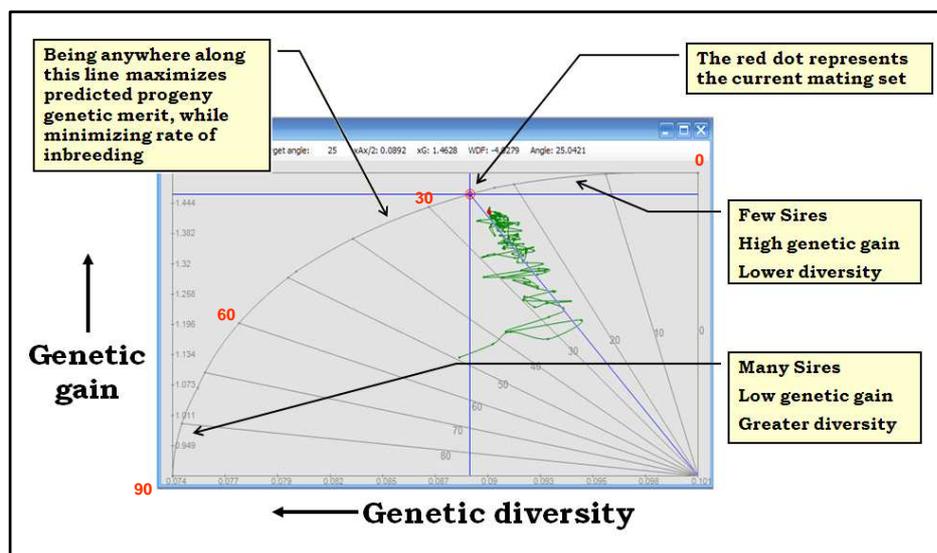


Figure 3.1 Graphical representation of the relationship between genetic diversity (x-axis) and predicted progeny gain (y-axis); target degrees in red figures (Newman et al., 2009, with graphic drawn using software by C. Gondro).

Where other component criteria are included in the objective function, such as progeny inbreeding, the frontier point is generally not reached.

However, the software manages the outcome such that the optimal solution will lie close to the target number of degrees line (Newman et al., 2009).

We used the following objective function:

$$\text{Obj} = x'G + \lambda_1 x'Ax + \lambda_2 F \quad (3.1)$$

where

- $x$  is the vector of selected parent or candidates contributions;
- $G$  is the vector of genetic merit of selected parents;
- $\lambda_1$  is the weight on coancestry;
- $A$  is the numerator relationship matrix among selected parents;
- $\lambda_2$  is the weight on short-term inbreeding;
- $F$  is the short-term inbreeding.

so

$x'G$  accounts for the predicted genetic merit in progeny in the next generation,  
 $x'Ax$  accounts for the long-term inbreeding as reflected by parental coancestry  
and  $F$  accounts for the predicted progeny mean inbreeding.

As previously mentioned, these simple cases were analysed:

- **a. Maximization of progeny merit**

In this case no constraints were imposed on inbreeding or coancestry. In Matesel target degrees were set to zero (equivalent to setting  $\lambda_1$  to zero) in order to maximize the predicted merit in the next generation.

- **b & c. Minimization of long-term or short-term inbreeding**

In order to show better how long-term and short-term inbreeding interact, cases b & c were analysed together.

In both cases Balance Strategy was set to 3. Ninety target degrees was chosen to obtain minimum coancestry in case 'b' (equivalent to setting  $\lambda_1$  to  $-999999$ ) and zero target degrees with  $\lambda_2$  being  $-999999$  to obtain minimum short-term inbreeding in case 'c'. Both scenarios 'b' and 'c' were repeated with 25 target degrees as that is usually sensible as it makes the sort of population structure that a practitioner might adopt (Kinghorn B., personal communication). In this way, reduction of coancestry was invoked in the case 'b' while in the case 'c' a negative weighting of -20 on progeny inbreeding was imposed to achieve reduced inbreeding in the next generation.

- **d. Variance of merit in progeny**

The criterion function in Matesel was modified in order to reward increased variance of predicted genetic merit (mid-parent EBV) of progeny in the next generation. So, it is necessary to previously calculate the variance of predicted genetic merit of progeny in the next generation for the mating list obtained in each run in order to reward in the criterion function that this variance has increased.

As in the cases b & c, Balance Strategy was set to 3 and target degrees were set to 25. Case 'b' with 25 target degrees was used as a reference scenario for comparisons (zero weighting on the progeny variance). Then, taking as a base the criterion function for case 'b' with 25 target degrees, two different runs were made: in the first one a small positive weight was imposed on the progeny variance and in the second a small negative weight was applied. The objective was to make no change in the selected parents, but a change in the pattern of mate allocations.

- **e. Look Ahead Mate Selection**

As previously mentioned, in this case the Look-Ahead Mate Selection scheme (LAMS) was applied and compared to the traditional scheme (no-LAMS) that aims to increase the merit of progeny in the next generation. The selection differential in the grand progeny generation was predicted by the difference between the average merit of *their* selected parents and the average merit of the entire population of current candidates before they are selected. To do this, the progeny predicted to be born from the current selected candidates were themselves put through a selection process, and the means of these selected progeny used to predict mean grand progeny merit.

Two different settings were tested: in the first, target degrees were set to zero while in the second were set to 45. In this way, it is possible to observe how the results were affected by coancestry. Balance Strategy was set to 3 in both cases. No constraints or weighting was imposed on inbreeding.

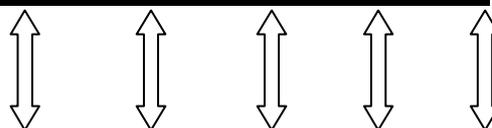
### ***3.4 Results and Discussion***

It is observed that giving more or less weighting or emphasis on one of the parameters genetic gain, short-term inbreeding or coancestry caused an impact on the other two, as expected.

The following table is presented showing the results for the different analysed cases.

Table 3.1 Results for the different analysed cases

Case	a	b	c	b	c	d		
Description	Max of Prog. Merit	Minim. of Coanc. 90 deg	Minim. of F 0 deg	Reduc. of Coanc. 25 deg	Reduc. of F 25 deg	- Wt on Var	No Wt on Var	+ Wt on Var
Selected Sires	10	21	15	13	14	13	13	13
Achieved Coancestry	0.0887	0.0381	0.0514	0.0675	0.0683	0.0675	0.0675	0.0675
Achieved Mean Progeny Index	56.957	47.719	48.302	56.038	55.856	56.038	56.038	56.038
Achieved Std Dev Progeny Index	3.662	3.924	4.487	4.214	4.482	3.129	4.214	5.193
Weighting on Progeny Mean F	0	0	-999999	0	-20	0	0	0
Achieved Progeny Mean F	0.073	0.027	0.018	0.052	0.037	0.049	0.052	0.057



Sire ID	Uses				
1064	19	25	20	19	20
1754	-	1	-	-	-
1775	25	23	24	25	24
1900	25	25	25	25	25
1959	23	20	24	23	24
2170	23	21	24	23	24
2351	25	25	25	25	25
2378	25	23	25	25	25
2381	14	16	12	14	12
2426	16	21	16	16	16
2469	23	23	24	23	24
2554	2	1	1	2	1
2590	5	1	5	5	5
2645	25	25	25	25	25

- **a. Maximization of progeny merit**

The following graphic is also presented showing other aspects of the resulting impact on the parameters in this case.

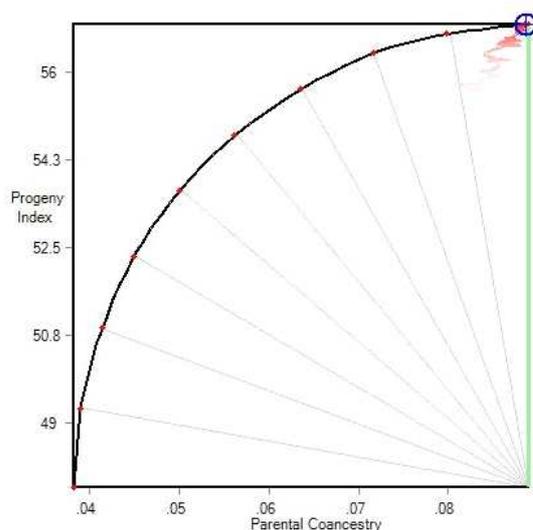


Figure 3.2. Maximization of progeny merit

Maximising genetic gain (see first column, Table 3.1 above) (in yellow colour) without consideration of the impact on long-term inbreeding resulted in high levels of coancestry (relatedness of individuals) (in green colour). In addition, in association with these genetic parameters, there was an increase of short-term inbreeding (in pink colour) and a reduction of the variance of the progeny merit in the next generation was obtained (in light blue colour). These results are observed not only in the corresponding calculated levels but also deduced as a consequence of the reduction in the number of the selected sires as parents (and consequently higher relatedness in their progeny) (in red colour). It is also observed that those sires with the highest EBVs, as well as the dams, were selected as parents.

In addition, other genetic parameters like the maximum F achieved, proportion of matings over the maximum F and the maximum F component criterion have varied accordingly (results not shown).

- **b & c. Minimization of long-term or short-term inbreeding**

In addition to the previous Table 3.1, the following graphics show other aspects of the resulting impact on the parameters in this case.

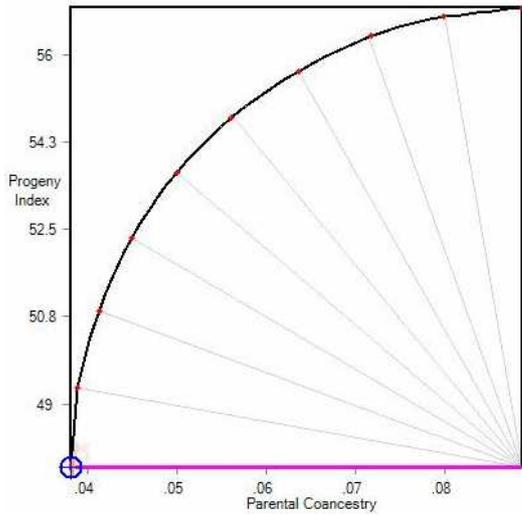


Figure 3.3. Minimization of Coancestry  
90 target degrees

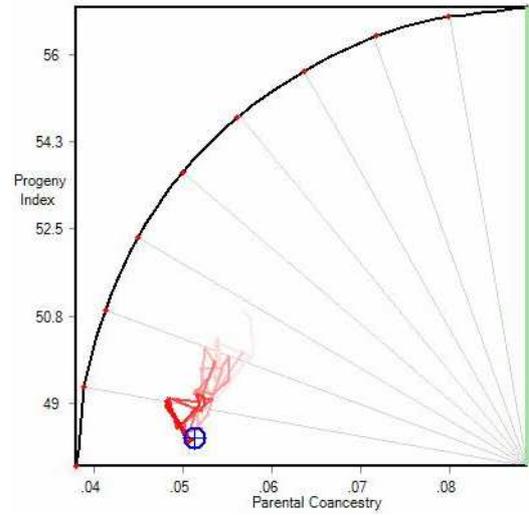


Figure 3.4. Minimization of  
progeny inbreeding (F)  
0 target degrees

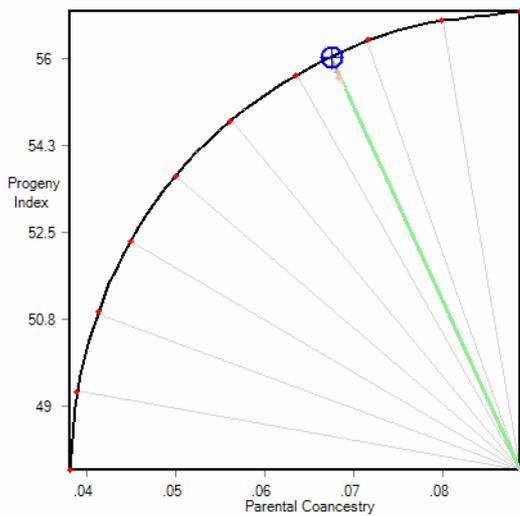


Figure 3.5. Reduction of Coancestry  
25 target degrees

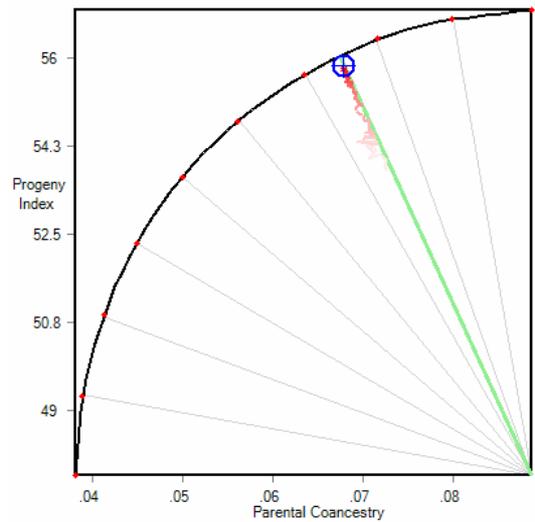


Figure 3.6. Reduction of  
progeny inbreeding (F)  
25 target degrees

Targeting 90 degrees, without consideration of the impact on other factors (see Table 3.1 above, second column of results), leads to the minimization of coancestry (increment of genetic diversity) (in green colour) and a significant reduction in the progeny merit in the next generation (in yellow colour). At the

same time, associated with these changes, a reduction in short-term inbreeding was observed (in pink colour).

On the other hand, the use of a very negative weighting for short-term inbreeding leads to its minimization (see Table 3.1 above, third column of results compared to the second column) (in pink colour) and a reduction in the number of the selected sires as parents (and consequently higher relatedness in their progeny) (in red colour). In addition, a small increase of coancestry (in green colour) and in the progeny merit in the next generation (in yellow colour) were observed.

The lower emphasis to reduce parental coancestry in the cases in which 25 target degrees were imposed compared to that in which 90 target degrees was imposed (see Table 3.1 above, fourth column of results compared to the second column) caused a lower impact on the behaviour of the parameters, but with similar trends. An increment of short-term inbreeding (in pink colour) and the progeny merit in the next generation (in yellow colour) were observed. The mentioned increment of short-term inbreeding (and, in consequence, the reduction of genetic diversity) was evidenced by the reduction in the number of the selected sires as parents (and consequently higher relatedness in their progeny) (in red colour).

The lower negative weighting used for short-term inbreeding (see Table 3.1 above, fifth column of results compared to the third column) leads to a small increase in coancestry (in green colour) and a small reduction in the progeny merit in the next generation (in yellow colour).

In addition, in both cases b & c, the observed change of the variance of the progeny merit in the next generation (in light blue colour) was due mainly to a variation in the pattern of mate allocations without changing significantly the selected sires and dams as parents (see first and second column of the lower part of Table 3.1).

The previous results are confirmed in the figures shown for these cases.

Thus, the lower emphasis to reduce parental coancestry is observed when 25 target degrees were imposed has resulted in higher Progeny Index (progeny merit in the next generation) (see Figure 3.7 vs. Figure 3.5). As expected, this effect increased as the target degrees was lowered.

On the other hand, it is observed that the higher negative weighting used for short-term inbreeding has resulted not only in considerable compromise in lower Progeny Index (progeny merit in the next generation) but also a lower parental coancestry (see Figure 3.6 vs. Figure 3.8). For practical purposes a much lower emphasis on progeny inbreeding can give useful results for progeny inbreeding with much less impact on Progeny Index and Parental Coancestry.

- **d. Variance of merit in progeny**

The following graphic shows additional aspects of the resulting impact on the parameters in this case.

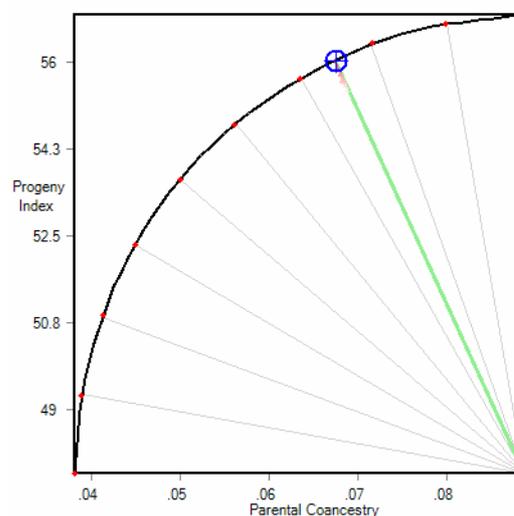


Figure 3.7. Variance of merit in progeny (no weight)

As previously mentioned, case ‘b’ with 25 target degrees was used as a reference scenario for comparisons (see upper part of Table 3.1 above, results in the fourth column are the same in the seventh column (zero weighting on the progeny variance)).

Comparing these results (see upper part of Table 3.1 above, sixth and eighth columns of results compared to the seventh column), an increase can be seen in the variance of the progeny merit in the next generation obtained by the use of positive assortative mating (in light blue colour).

In addition, the lower part of Table 3.1 shows the change caused in the pattern of mate allocations with only small changes of the number of matings for selected sires and dams as parents when small weightings (negative or positive) were given on the SD of the progeny merit.

Although there were no changes in other genetic parameters in Table 3.1 using these small weightings on the SD of the progeny merit it is shown that an increment of the variance of the progeny merit in the next generation is associated with an increase in short-term inbreeding (in pink colour).

- **e. Look Ahead Mate Selection**

The following Table 3.2 and graphics show other aspects of the resulting impact on the parameters in this case.

Table 3.2 Look Ahead Mate Selection

Item	Scheme	no-LAMS	LAMS	no-LAMS	LAMS
		(0 deg)	(0 deg)	(45 deg)	(45 deg)
Selected sires:		10	12	15	18
Achieved Coancestry:		0.089	0.085	0.053	0.058
Achieved Mean Progeny Index:		56.957	56.375	54.232	53.313
Achieved Std Dev Progeny Index:		3.712	4.881	5.019	6.035
Achieved Progeny Mean F:		0.073	0.066	0.046	0.050
Maximum F achieved:		0.313	0.281	0.281	0.281
Proportion of matings over MaxF:		0.040	0.012	0.028	0.024
MaxF component criterion:		7.930	2.392	5.824	4.420
Mean Progeny 2nd Generation		60.790	61.503	59.267	59.895

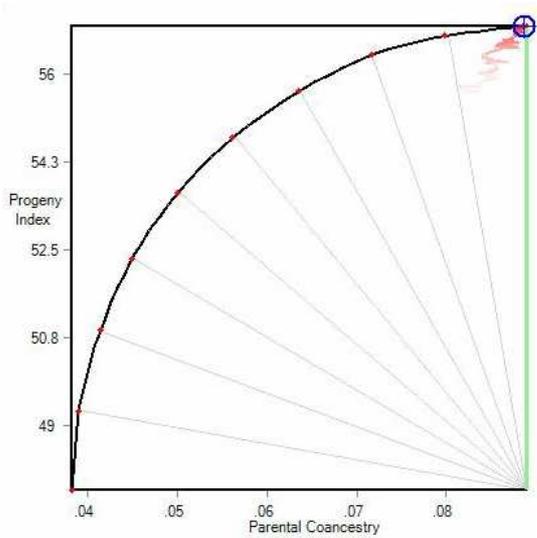


Figure 3.8. no-LAMS (0 deg)

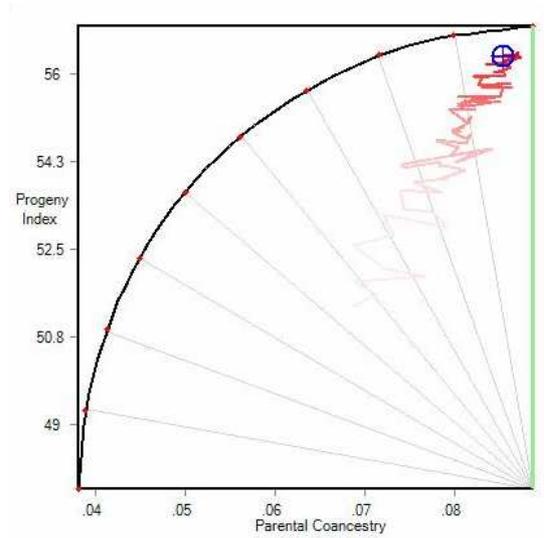


Figure 3.9. LAMS (0 deg)

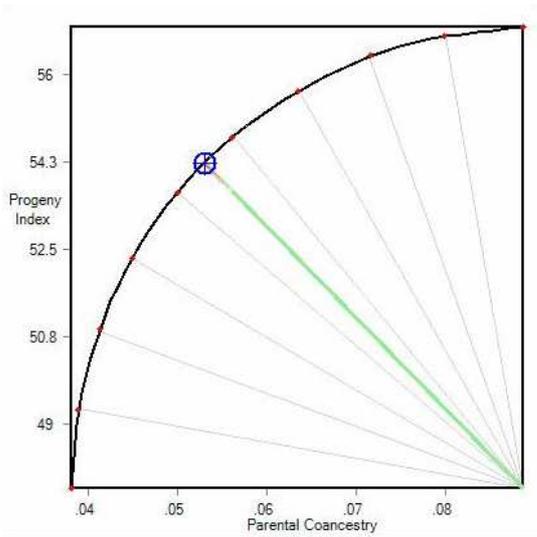


Figure 3.10. no-LAMS (45 deg)

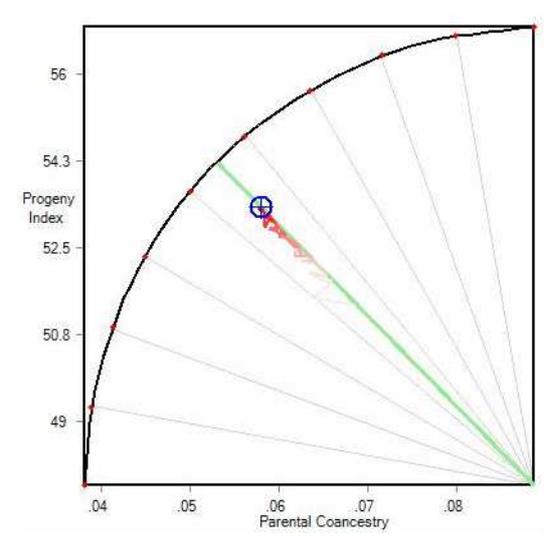


Figure 3.11. LAMS (45 deg)

As in the other cases, a reduction of the levels of coancestry (see above Table 3.2) (in green colour) corresponded to a reduction of the levels of short-term inbreeding (in pink colour), a reduction of the genetic gain in the next generation (in yellow colour), an increment of the variance of the progeny merit in the next generation (in light blue colour) and an increment in the number of the selected sires as parents (and consequently lower relatedness in their progeny) (in red colour).

It is observed that LAMS has resulted in some compromise in lower Progeny Index (progeny merit in the next generation) (see Figures 3.9 & 3.11 vs. Figures 3.8 & 3.10) and slightly lower parental coancestry at 0 degrees (see Figures 3.8 vs. 3.9) but slightly higher parental coancestry at 45 degrees (see Figures 3.10 vs. 3.11).

In addition, LAMS caused an increment of the variance of the progeny merit in the next generation (in light blue colour). This increment allowed bigger differential selections and, in consequence, bigger genetic gains in the second generation (in orange colour).

Although the increment of the variance of the progeny merit was bigger (1.169 vs. 1.016) as the coancestry was reduced (changing from 0 degrees to 45 degrees) (in light blue colour), in this case, however, relatively bigger genetic gains in the second generation were obtained (in orange colour) with higher levels of coancestry (in green colour) (0 degrees vs. 45 degrees). Effects on progeny inbreeding and coancestry were small and differed between 0 degrees and 45 degrees, with decreases for the former and increases for the latter.

### ***3.5 Conclusions***

Optimal balance between high genetic gain and low parental coancestry when selecting breeding animals can be achieved with the optimal set of contributions of the selected individuals' genes to the breeding program. These optimal contributions can be found by using a stochastic search algorithm, such as a Differential Evolution algorithm. Additional attention to mate allocation, using simultaneous Mate Selection, adds management of progeny inbreeding, progeny variance and predicted grand progeny merit, as tested here, plus other factors.

In all the analysed cases a reduction in the levels of coancestry corresponded to a reduction in genetic gain in the next generation. At the same

time these effects were associated with a reduction in the levels of short-term inbreeding, an increment of the variance of the progeny merit in the next generation and an increment in the number of the selected sires as parents (and consequently lower relatedness in their progeny).

The application of a Mate Selection tool can lead to sustainable genetic progress, with maintenance of genetic diversity and avoidance of the undesirable consequences of inbreeding. Adding other technical and logistical factors allows practical implementation of sustainable breeding programs.

LAMS approach has resulted in some compromise in slightly lower progeny merit in the next generation, but caused an increase in the variance of the progeny merit in the next generation. This increase allowed bigger differential selections and, in consequence, bigger genetic gains in the second generation, i.e. the predicted grand progeny merit. The results of this general test show some promise for the LAMS approach, and it will be further tested for specific scenarios on Chapters 5 and 6.

Although the results of the cases shown in this Chapter must be considered illustrative rather than evaluative, the link between Mate Selection method and results achieved is direct, and less subject to sampling error because contrasted treatments use the same dataset.

There seems to be little or no previous work similar to the analyses carried out in this Chapter, probably because of a relative lack of development of Mate Selection systems that operate beyond control of genetic merit and coancestry.

## **CHAPTER 4**

# **EXPLORING NOVEL BREEDING PROGRAM APPLICATIONS USING MATE SELECTION**

## ***4.1 Summary***

This chapter investigates the consequences of applying different tactical mate selection approaches in stud-farms with a moderate to large herd size that apply a two-tier nucleus breeding scheme. In this scheme the nucleus is closed to the lower tier but open to imported animals, semen or embryos from other nuclei. A real data base of leader Argentinean Angus purebred stud-farms is used for analysis.

In this Chapter, different tactical scenarios are analysed. The first shows the impact of the use of various reproductive techniques in a breeding program on predicted genetic merit of progeny and other genetic parameters according to a restriction on budget. The second scenario shows the application of Mate Selection for dissemination, i.e. dispersal of breeding males to different beef production units and the achievement of a bimodal population, both oriented to diverse beef markets.

The results show that Mate Selection allows the breeder to take decisions to balance the use of various reproductive technologies, and hence to choose the most suitable set of matings, in order to achieve his/her breeding objectives taking into account the restrictions on budget imposed. It is also shown that Mate Selection is suitable to simultaneously select different sets of sires in the stud-herds according to the particular breeding objectives of different commercial units, even including those of the stud. The consideration of these objectives in the MSI function allows the optimisation of profit over herds, in addition to the control of other genetic aspects such as maintenance of genetic diversity and avoidance of the undesirable consequences of inbreeding.

## ***4.2 Introduction***

Herd size is a factor that affects genetic response due to its relation to selection intensity and inbreeding, especially in the long-term, through effects on effective population size. Also the availability of more phenotypes probably leads to more accurate genetic evaluation of animals. According to this, two types of breeding enterprises of purebred beef cattle exist in Argentina. One corresponds to stud-farms with a limited herd size and the other to stud-farms with a moderate to large herd size. The latter applies a two-tier nucleus breeding scheme, in which the nucleus is closed to the lower tier but open to imported animals, semen or embryos from other nuclei. This chapter investigates the consequences of applying different tactical mate selection approaches in the mentioned scenario for stud-farms with a moderate to large herd size. This includes the use of diverse reproductive techniques and dissemination of breeding males to different beef markets considering different kinds of costs and constraints. A real data base of leader Argentinean Angus purebred stud-farms is used for analysis.

It should be mentioned that the approaches shown in the following analysed cases are only three out of a range of issues that could be managed in this way.

## ***4.3 Materials and methods***

### **4.3.1 General description of the data used in these analyses**

The data set used belongs to an Angus herd of a seed-stock farm sited in the humid temperate region of Argentina. This herd was previously genetically evaluated under the Argentine National Angus Genetic Evaluation, using a recorded pedigree.

Three economic indices for 3 beef markets in Argentina were constructed based on breeding objectives derived from guidelines set by Ávila, 2006 and Ávila et al., 2004 (see below) and have been complemented with concepts provided through personal communications by the breeders Bustillo P. and Lafontaine J. A., who are the owners of “La Llovizna” and “Los Tigres” respectively, which are two leader Angus purebred stud-farms sited in the temperate region of Argentina.

The markets involved are:

- a. Domestic market
- b. Certified Angus Beef market
- c. Export market

The mentioned guidelines for these beef markets from which their corresponding breeding objectives were derived are:

– Domestic market

- Calving ease (Avg or negative Birth Weight EBV)
- positive Weaning Weight EBV
- positive Final Weight EBV (18-months age): Endpoint 360-420 Kg
- Frame Score: 3-4
- positive Ribeye Area EBV
- Back Fat Thickness: 4 to 8 mm

– Certified Angus Beef market

- Calving ease (Avg or negative Birth Weight EBV)
- positive Weaning Weight EBV
- positive Final Weight EBV (18-months age): Endpoint 410-470 Kg
- Frame Score: 4-5
- positive Ribeye Area EBV
- positive %IMF EBV
- Back Fat Thickness: 4 to 8 mm

– Export market

- Calving ease (Avg or negative Birth Weight EBV)
- positive Weaning Weight EBV
- positive Final Weight EBV (18-months age): Endpoint 460-520 Kg
- Frame Score: 5-6
- positive Ribeye Area EBV
- positive %IMF EBV
- Back Fat Thickness: 5 to 12 mm

Taking into account these guidelines, a suitable group of evaluated traits to build the three mentioned indices were considered. Calving rate is not currently addressed.

Actual regional costs of beef production, artificial insemination centres and market prices were used to calculate the economic weights of the considered traits. The economic weights were expressed in dollars per head.

In our case, this other information was also taken into account:

- Variables

- Market price or income of the considered traits in dollars per head,
  - Costs or expenses of the considered traits in dollars per head,

- Constants

- Market or sale weight according to the specific index,
  - Days to Slaughter.

Then, using all this information, a production model was built up, in which a suitable average value was assigned to each of the considered traits and the other variables and constants according to the index that was been analysed. Thus, finally, subtracting the total calculated costs or expenses from the total calculated income, the **net income or profit in dollars per head** of our production model was obtained.

Once the production model was built up, knowing the net income or profit of it using the average values for traits and the other variables and constants

involved, one of the traits in one unit (expressed in dollars per unit) was increased, keeping the other traits constant. This caused the change of the net income or profit of the model. At last, subtracting this new resulting net income from the previous one, the economic weight of the trait (expressed in dollars per unit) that had been increased in one unit was calculated. These last steps were repeated for all the considered traits of the index in order to estimate their corresponding economic weight.

The genetic parameters that are presented in the following Table 4.1 were used for the calculation of the economic index:

<b>Table 4.1 Heritability and phenotypic variance for the considered traits for indices</b>							
<b>Trait</b>	<b>BW</b>	<b>WW</b>	<b>18m WT</b>	<b>18m HT</b>	<b>REA</b>	<b>RF</b>	<b>%IMF</b>
Heritability	0.30	0.25	0.30	0.47	0.36	0.41	0.36
Phenotypic Variance	16.2	810.0	1280.0	17.0	148.9	4.8	2.2

where BW refers to Birth weight, WW to Weaning weight direct, 18mWT to 18-Month weight, 18mHT to 18-Month height, REA to Ribeye area, RF to Rump fat thickness and %IMF to % Intramuscular fat.

The calculated economic weights for these traits according each one of the three selection indices are presented in the following Table 4.2:

<b>Table 4.2 Economic weights for the considered traits for indices</b>							
<b>Trait</b>	<b>BW</b>	<b>WW</b>	<b>18m WT</b>	<b>18m HT</b>	<b>REA</b>	<b>RF</b>	<b>%IMF</b>
<b>Domestic market</b>	-8.135	0.792	0.618	-2.175	0.700	2.922	0.000
<b>Certified Angus market</b>	-6.127	0.648	0.523	-2.357	0.770	0.000	6.333
<b>Export market</b>	-6.687	0.802	0.609	-2.532	0.675	0.000	10.439

Once the economic weights have been estimated, the relative economic weight for each trait was calculated. For this, the products of each economic weight by their trait's estimated genetic standard deviation were computed (equation [4.1]). These relative economic weights can be expressed as an emphasis or economic importance for each EBV in percentage values by dividing each relative economic weight by the sum of the absolute relative economic weights of the considered traits for an index and multiplying each result by 100 (equation [4.2]).

$$rev_i = v_i * SDa_i \quad [4.1]$$

where *rev* refers to the relative economic weight of trait *i*, *v* to the economic weight of trait *i* and *SDa* to the estimated genetic standard deviation of trait *i*.

$$\% \textit{emphasis}_i = (rev_i / \sum_{i=1}^{trait} |rev_i|) * 100 \quad [4.2]$$

where *%emphasis* refers to the percentage emphasis or economic importance for trait *i*.

Finally, the economic index for each animal was computed by summing the products of those calculated relative economic weights by their corresponding EBV, the latter being previously divided by their trait's estimated genetic standard deviation (equation [4.3]).

$$I = rev_1 (EBV_1 / SDa_1) + rev_2 (EBV_2 / SDa_2) + \dots + rev_n (EBV_n / SDa_n) \quad [4.3]$$

where *I* refers to the economic index.

The percentage emphasis for each EBV of the considered evaluated traits in the Domestic, Certified Angus Beef and Export market indices are shown in Figure 4.1.

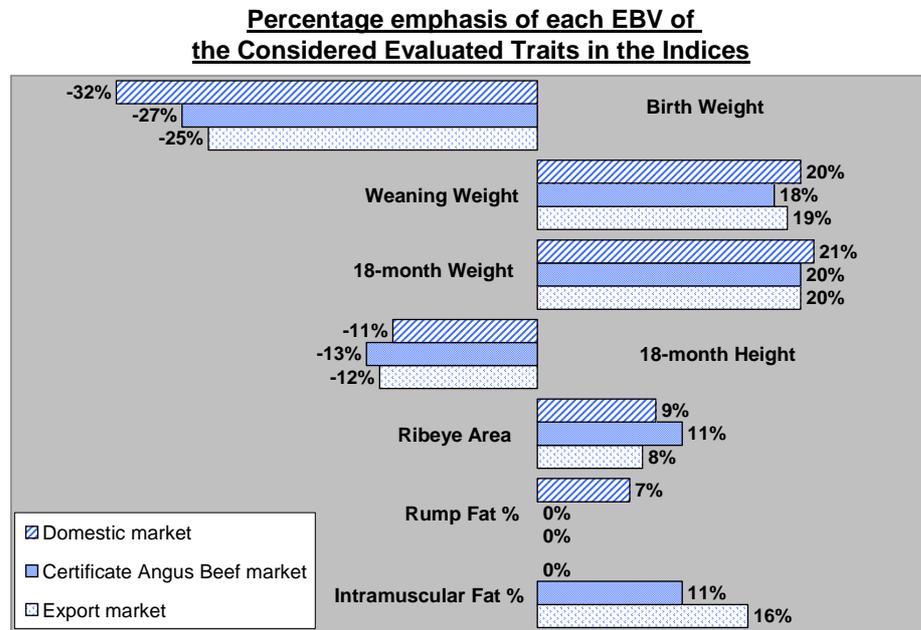


Figure 4.1 Relative emphasis of each EBV of the considered evaluated traits in the Domestic, Certified Angus Beef and Export market indices.

Knowing the economic importance of our traits, the economic weights and the breeding values of animals for selection decisions can be used.

Thus, these economic indices respond to suitable breeding objectives according the requirements of those beef markets.

### 4.3.2 General analysis method

According to the considered tactical approach, the Matesel option in the Pedigree Viewer system (<http://www-personal.une.edu.au/~bkinghor/pedigree.htm>) was then applied to determine the

set of matings (including optimum contribution of candidates) to generate the next set of progeny taking into account the objectives of each case. The source code of the Matesel system was modified by the author in order to achieve the custom analyses required in this chapter.

As previously described in Chapter 3, the Matesel system can balance progeny index response (genetic gain), parental coancestry and other factors, like costs, in several ways. Thus, undesired solutions are penalised, considering Balance Strategy settings, different weightings (e.g. on coancestry, on progeny inbreeding, etc.) and the other factors considered in the Mate Selection Index (MSI) function.

### **4.3.3 The analysed cases**

The different tactical approaches to be analysed in this Chapter were:

- Case A) Impact of the use of various reproductive techniques in a breeding program on predicted genetic merit of progeny and other genetic parameters according to a restriction on budget.
- Case B) Dissemination: Dispersal of breeding males to different beef production units oriented to diverse beef markets.
- Case C) Mate Selection of breeding males and females to obtain different proportions of the herd oriented to two different beef markets.

#### **4.3.3.1 Case A) Impact of the use of various reproductive techniques in a breeding program on predicted genetic merit of progeny and other genetic parameters according to a restriction on budget**

#### **4.3.3.1.1 Objectives of this analysis**

The objectives of the analysis were:

- to illustrate the pattern of change in the number of Natural, AI and MOET matings as the budget increases;
- to investigate the impact of the number of Natural, AI and MOET matings on genetic merit, the variance of genetic merit and parental coancestry;
- to investigate the impact of costs associated with the use of different reproductive techniques on the different aspects of response.

#### **4.3.3.1.2 Data description**

The data set of the analysed herd consisted of 905 female candidates divided into 2 groups (151 purebred females (the nucleus) and 754 multiplier females) and 47 Purebred male candidates (24 potential AI and 23 natural mating sires). The relatively small number of male candidates makes sense due to higher prior selection intensity applied to males than to females. Numbers of male candidates is relatively low due to pre-selection. Each candidate had its corresponding pedigree record. Genetic merit was expressed in Certified Angus Beef index units.

The mean index value of sires for AI was higher than the mean index value of sires for natural mating, which, at the same time, was higher than female candidates (Table 4.3). Again, this makes sense due to higher prior selection among males.

Table 4.3: Statistics of candidate animals in Certified Angus Beef index units

Candidate animals (Genetic merit expressed in Certified Angus Beef index units)				
	NatMat Sires	Dams (multiplier)	AI Sires	Dams (purebred)
N	23	754	24	151
Mean Index	33.4	29.5	37.2	31.1
Index StdDev	5.7	6.2	6.6	5.8

Nominated maximum number of matings or planned pregnancies for each candidate was considered. In this way, up to 40 matings were assigned for natural mating sires, 1000 for AI sires, up to 5 for females that can enter in a MOET program and only 1 (or zero) for females that cannot. Here, it is convenient to take into account the difference between MOET (Multiple Ovulation and Embryo Transfer) and JIVET (Juvenile in Vitro fertilization and Embryo Transfer). In the former technique, the super-ovulated female is fertilized by a single sire, so all the obtained embryos come from the same sire and dam. In the latter, not necessarily all the picked oocytes from a young female are in vitro fertilized with semen of the same sire. This fact makes a difference in parameterization for a Mate Selection analysis.

In addition, the cost of using each sire was included, taking into account the reproductive technique for which it would be used.

An input file was built using the previously mentioned data set.

Costs and benefits were expressed in dollars per calf born.

- **Costs:**

The costs presented and used in this case were obtained through personal communications with the previously mentioned breeders of advanced seed-stock

farms of the temperate region of Argentina plus summaries of final prices of auctions of purebred Angus livestock and price lists of different centres of embryo transfer and artificial insemination of this region.

The cost of a natural mating was calculated as the average price of a replacement purebred sire divided by the average effective number of progeny in its average lifetime.

This is:

avg price of a purebred sire = \$ 8600

avg effective number of progeny per year = 35 to 40 (average 38)

avg lifetime of a sire = 3 years

It was considered that the cost of maintenance of the sire (feeding, veterinary costs, etc.) is completely recovered at the moment of culling when it is sold to the slaughterhouse, so this cost was 0 in the analysis.

So

cost of a natural mating =  $8600 / (38 \times 3) \approx \$ 75$

On the other hand, to calculate the cost of an AI mating it should be considered the operational cost per mating (drugs, veterinary costs, etc.), the cost per dose of semen and the percentage of failure.

This is:

operational cost per female = \$ 105

cost per dose of semen = \$ 5 to 16 (average \$ 11)

percentage of failure = 50%

So

cost of an AI mating =  $105 + (11 \times 2) \approx \$ 115$  to 137  
(depending on the cost per dose of semen)

Finally, to calculate the cost of MOET it was considered that 5 embryos in average per round would result in birth per female included. It was calculated that the total cost per female per round of MOET is \$ 236 per calf born plus the proportional cost of at least 2 doses of semen. This total cost comprises the cost of drugs, veterinary costs, synchronization of recipients and cost of implantation.

The total reproductive cost or budget was expressed as the average cost per mating, i.e. the average cost in dollars per calf born.

#### **4.3.3.1.3 Analysis method**

As in previous analyses, the Matesel option in the Pedigree Viewer system was applied to determine the set of matings to generate the next set of progeny.

Balance Strategy was set to 3, which put a hard constraint on the selected number of target degrees, severely penalising solutions that break the target degrees constraint. In this scenario no constraints were imposed on progeny inbreeding. In Matesel target degrees were set to 25, imposing a moderate constraint on parental coancestry (see previous section 3.3.3 to refresh the description of these Matesel parameters, if necessary).

The way to carry out the analysis was to progressively increase the available total budget for reproduction and, in consequence, the possibility of using a higher proportion of more expensive reproductive techniques. This means that the dollar value of index values did not have to be aligned with the dollar value of reproductive costs, which is difficult in practice. It should be noted that there is no simultaneous optimisation of animal selection and money spent on reproduction, as the value of these two components is not easy to compare. This leads to a desired outcomes solution based on the range of results achieved. Thus, according to the available total budget for reproduction, a resulting mating list is obtained through Mate Selection which maximises the

objective function, establishing in this way the proportion of mates for each of the used reproductive technologies. In addition, the impact of different genetic parameters was analysed.

#### **4.3.3.2 Case B) Dissemination. Dispersal of breeding males to different beef production units oriented to diverse beef markets**

Using the previously mentioned Angus data set from an Argentinean seed-stock farm, a mate selection analysis was run to select sires into the stud for two different beef markets. So, dispersal of sires to different breeding units is achieved, i.e. oriented to two different beef markets, handled simultaneously with selection of sires into the stud.

The herd was evaluated using economic indices described above, built according the requirements of the following markets:

- a. Domestic market,
- b. Export market.

Thus, the Purebred sires of the nucleus were used on the females of the nucleus and on the complete Multiplier herd of cows in order to get progeny with the optimum genetic mean in the Export market index in the former and in the Domestic market index in the latter.

For comparison, a different Mate Selection analysis was run in which an average of the Export market and Domestic market indices in both herds was used, i.e. the Purebred sires of the nucleus and the females of both herds were selected and mated to get progeny with the optimum genetic mean in this average index.

As both herds of cows had a recorded pedigree, additional genetic objectives were achieved using suitable constraints on coancestry and inbreeding.

Finally, the results of these two strategies were compared.

Dissemination cases like these can be managed by running analyses in which grouping constraints are applied on certain matings, as dictated through a group mating permission matrix, as described by Kinghorn (2011), where a sire can be used in some herds but not in the others. In this way, for this case, Multiplier sires were not permitted to be mated while Purebred sires could be mated to females of the nucleus and of Multiplier herd. Also, the code of the Mate Selection program was modified to indicate which kind of index should be used with which group in order to allocate matings to markets, according to which of the two strategies for this case was carried out.

#### **4.3.3.2.1 Objectives of the analysis**

The objectives of the analysis were:

- to show using real data how dispersal of sires to different breeding units can be handled simultaneously with selection of sires into the stud,
- to show that the different sets of sires selected for diverse breeding objectives can move the progeny index mean in the desired directions for each herd,
- to show that this approach can simultaneously optimise other genetic parameters and profit across herds.

#### **4.3.3.2.2 Data description**

Based on the data set of a seed-stock herd in Argentina, described previously, a part of it was analysed that consisted of 24 Purebred sire candidates from the nucleus and 905 female candidates divided in two groups: Purebred (the nucleus) and Multiplier. As previously mentioned, all animals in

the herd had their corresponding pedigree record and genetic merit expressed in units of the considered indices.

#### **4.3.3.2.3 Analysis method**

As for the previous analysis, the Matesel option in the Pedigree Viewer system was applied to determine the set of matings to generate the next set of progeny.

Balance Strategy was set to 3, which put a hard constraint on the selected number of target degrees, severely penalising solutions that break the target degrees constraint.

Firstly, dissemination was analysed using 25 target degrees to put a moderate emphasis on reduced parental coancestry. Then, the effects of targeting low coancestry alone and strong weighting against inbreeding were investigated on their impact on genetic diversity and profit. The settings used for this were 0 target degrees and weighting coefficient  $-999999$  on progeny inbreeding level.

In order to do fair comparisons between the results of both strategies, each mating in the sets of matings obtained when the indices were used separately was identified to know which of them were destined for Export market and which for Domestic market. Then, in the set of matings obtained when the average index was used, the average index of the sire and the female of each mating was replaced with their corresponding Export market or Domestic market index according to which market that female was destined in the same scenario of the other strategy. Finally, the achieved mean progeny indices were recalculated.

#### **4.3.3.3 Case C) Mate Selection of breeding males and females to obtain different proportions of the herd oriented to two beef markets.**

Using the previously mentioned Angus data set from an Argentinean seed-stock farm, a mate selection analysis was run to obtain a bimodal population, i.e. different proportions of the herd oriented to two different beef markets.

As indicated, the herd was evaluated using economic indices built according the requirements of these markets. The economic indices considered this analysis were:

- a. Domestic market,
- b. Export market.

Thus, the proportion of the two groups of the herd, each one focused to one of the two indices, can be set equal or to specified proportions.

Then, a new Mate Selection analysis based on an alternative strategy was run which used an average of the Export market and Domestic market indices, i.e. a set of matings was obtained to get progeny with the optimum genetic mean in this average index.

Finally, the results of these two strategies were compared.

As in the previous case B, these analyses used the dissemination concept, as described by Kinghorn (2011), in which grouping constraints were applied on certain matings, as dictated through a group mating permission matrix, where matings were not permitted between certain groups. So, for this case, Multiplier sires and females were not permitted to be mated while Purebred sires could only be mated to females of the nucleus. In addition, the code of the Mate Selection program was modified to indicate which kind of index should be used with which group or proportion of the herd in order to allocate matings to markets, according to which of the two strategies for this case was carried out.

#### **4.3.3.3.1 Objectives of the analysis**

The objectives of the analysis were:

- to show using real data how selection and allocation of sires and dams can be handled to different breeding objectives,
- to show that the different sets of sires and dams selected for diverse breeding objectives can move the progeny index mean in the desired directions.

#### **4.3.3.3.2 Data description**

Differently to the previous case B, this analysis used the data set of the nucleus (Purebred group) of the seed-stock herd that consisted of 151 female and 24 Purebred male candidates. As previously mentioned, each candidate had its corresponding pedigree record and genetic merit expressed in units of the two indices considered.

#### **4.3.3.3.3 Analysis method**

As in the previous analyses, the Matesel option in the Pedigree Viewer system was applied to determine the set of matings to generate the next set of progeny.

Firstly, Balance Strategy was set to 3, which put a hard constraint on the selected number of target degrees, severely penalising solutions that break the target degrees constraint.

In addition, dissemination was analysed using 25 target degrees to put a moderate emphasis on reduced parental coancestry. The proportion of the herd focused at each of the two indices was varied. Three proportions were chosen: a) 0.2:0.8, b) 0.5:0.5 and c) 0.8:0.2.

Then, using the same target degrees and Balance Strategy, a different Mate Selection analysis was run in which the Purebred sires and females of the nucleus were selected and mated to generate progeny with the optimum genetic mean using an average of the Export market and Domestic market indices.

As in the previous case 'B', in order to do fair comparisons between the results of both strategies, each mating in the sets of matings obtained when the indices were used separately was identified to know which of them were destined for Export market and which for Domestic market. Then, in the set of matings obtained when the average index was used, the average index of the sire and the female of each mating was replaced with their corresponding Export market or Domestic market index according to which market that female was destined in the same scenario of the other strategy. Finally, the achieved mean progeny indices were recalculated.

#### ***4.4 Results and Conclusions***

##### **4.4.1 Case A) Impact of the use of various reproductive techniques in a breeding program on predicted genetic merit of progeny and other genetic parameters according to a restriction on budget**

The results obtained showed that the total reproductive cost increased as AI and/or MOET matings were included in the program.

On the other hand, the number of Natural matings decreased as AI matings increased. This trend also appears as the number of MOET matings increased but with lower number of Natural and AI matings (see Figure 4.2).

At the same time, it was observed that progeny merit in the next generation increased (see Figure 4.3) as the total reproductive cost increased.

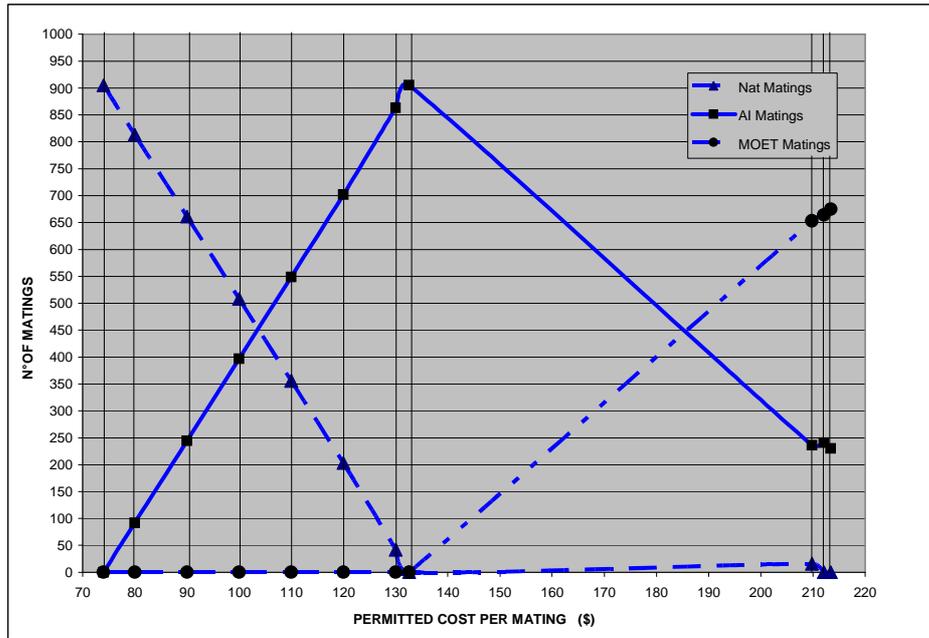


Figure 4.2 Relationship between No. of Natural, AI and MOET matings and restriction on budget imposed. The total number of matings at a specific cost is equal to the sum of matings of each reproductive technique along the vertical line coincident to that cost.

It should be noted that the trend in Figure 4.3 was due to progressively higher average costs due to increasing use of AI and MOET matings respectively together with the higher genetic merit of AI sires and MOET sires and dams compared to those used for Natural matings.

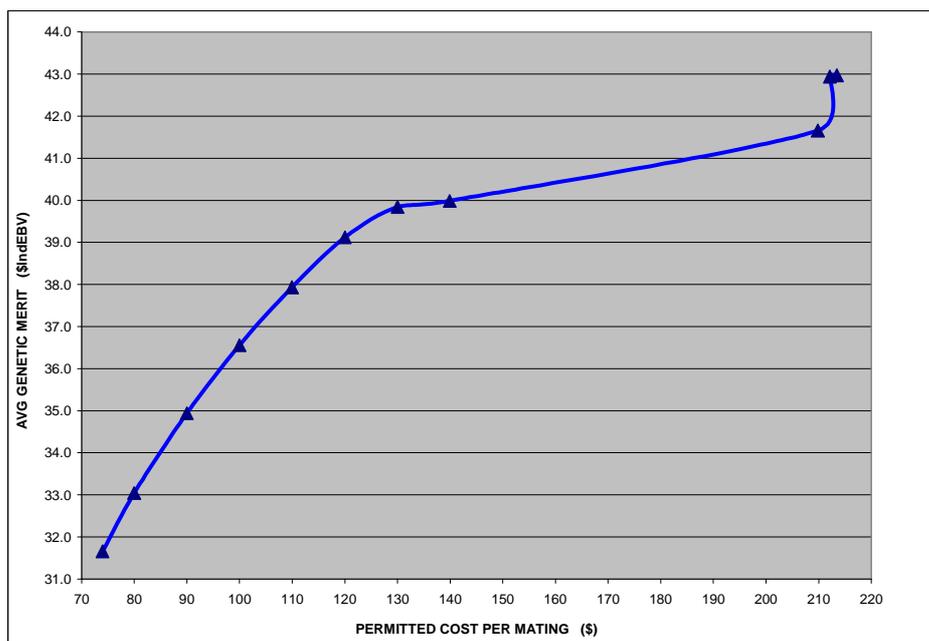


Figure 4.3 Relationship between predicted genetic merit of progeny and restriction on budget imposed.

Note also a sudden very significant increase in the genetic merit, despite a slight increase of the average cost per mating, at the end of the trend in Figure 4.3. This is caused by the start of the use of MOET. So, only after the budget allows an average cost per mating equal to \$210, MOET can be used. This means that the number of AI matings remains almost the same from an average cost per mating equal to \$140 to another equal to \$200 (see also Figure 4.3).

At the same time, it is observed that, even though there was a higher predicted genetic merit of progeny as the number of AI and MOET matings increased, the ratio \$EBV/cost per mating is higher for AI than for MOET matings (see Figure 4.3), i.e. the use of AI caused a proportional higher increment in the predicted genetic merit of progeny than MOET (see Figure 4.4). Thus, the relatively high cost of MOET matings would make the use of this reproductive technique only worthwhile for progeny of high predicted genetic merit.

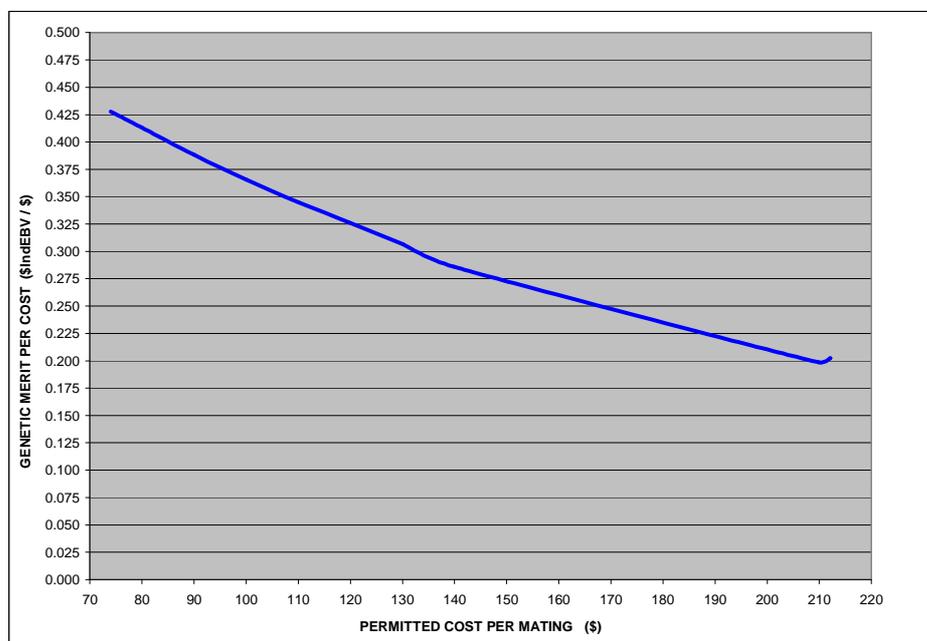


Figure 4.4 Relationship between \$EBV of predicted genetic merit of progeny per dollar of cost of mating and restriction on budget imposed.

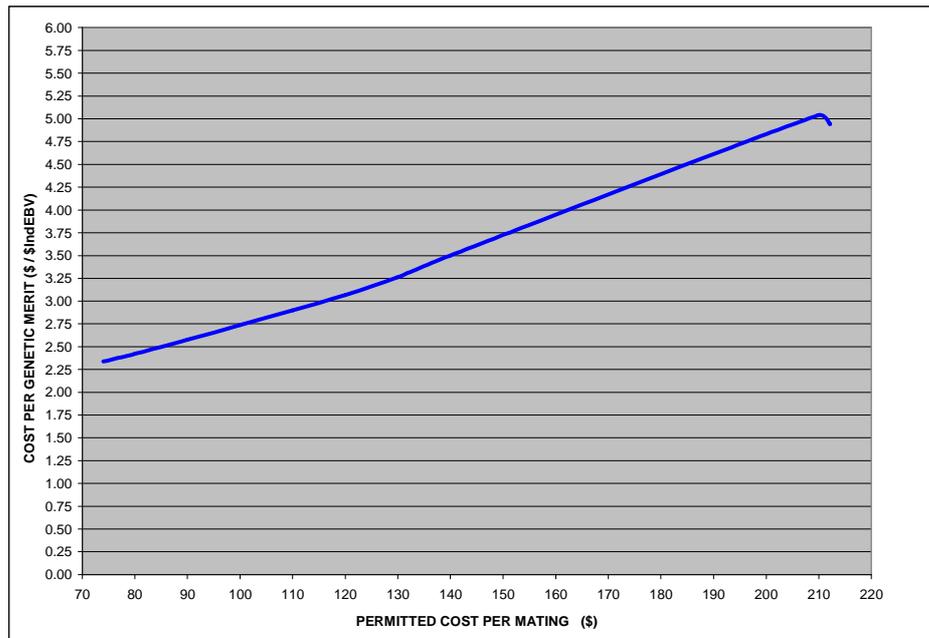


Figure 4.5 Relationship between cost per \$EBV of predicted genetic merit of progeny and restriction on budget imposed.

On the other hand, a positive relationship between coancestry and cost per mating is seen in Figure 4.6, i.e. coancestry increased as the average cost per mating increased. This was caused by the progressive reduction in the number of the selected sires as parents for Natural and AI matings and the higher number of matings assigned mainly to the selected sires and dams for MOET (and hence higher relatedness in their progeny) as the restriction on budget imposed decreased. Note that differences in coancestry can result because a fixed limit on coancestry was not applied. The 25 degree constraint is effectively different for each cost level, as a balance between genetic gain and coancestry is made according to the reproductive opportunities available. A hard fixed constraint might result in no use of MOET even at high permitted cost. With more reproductive opportunities, both coancestry and gain will generally be higher under the most aggressive policy (zero degrees or maximal high gain). Less aggressive policies reduce both gain and coancestry at each level of reproductive cost, as seen in Figure 4.7.

The trends of the relationships genetic merit of progeny vs. coancestry (Figure 4.7) and coancestry vs. cost per mating (Figure 4.6) were a consequence

of the higher number of matings assigned to the selected sires with higher \$EBVs as the permitted budget increased. At the same time, it is observed that the increment of genetic gain and coancestry become smaller as the permitted budget increased (see the top right of Figures 4.3 and 4.5).

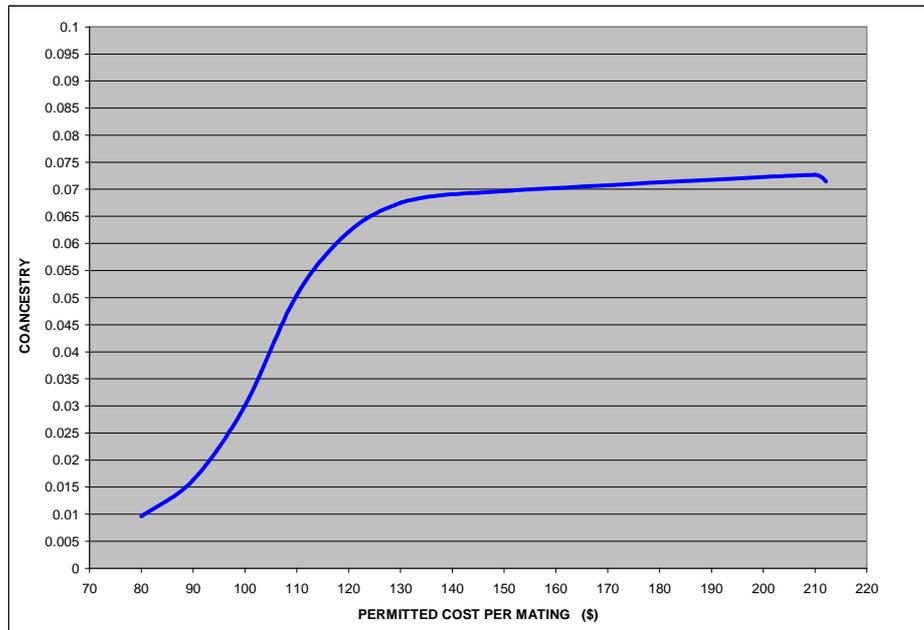


Figure 4.6 Relationship between coancestry and restriction on budget imposed

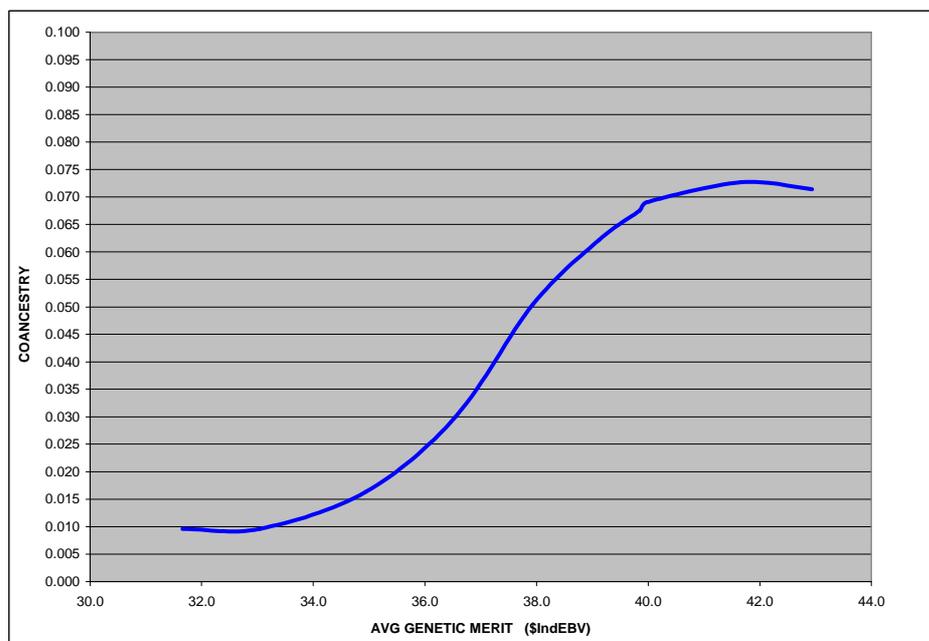


Figure 4.7 Relationship between predicted genetic merit of progeny and coancestry due to the higher number of matings assigned to the selected sires with higher \$EBVs as the permitted budget increased

Analysing the final selected sire candidates, it is noted that in order to select a sire among the possible candidates the Matesel program considered not only the cost per mating that each sire generates but also at the same time the genetic merit of each sire, and its impact on average parental coancestry.

#### **4.4.1.1 Conclusions**

This section investigated the impact of budget restrictions on use of different reproductive techniques, and therefore the set of selected sires and dams. Simultaneously with the variation of the number of Natural, AI and MOET matings, the components of response varied, including the predicted genetic gain in progeny and coancestry. So, reproductive technologies were seen at the same time from a genetic and economic point of view.

All these factors, including use of reproductive technologies, can be optimised by including them in the objective function. Applying a constraint on funds used for reproductive technologies can be used as appropriate.

Using this tactical approach, the breeder can take decisions to balance the use of these reproductive technologies, and hence to choose the most suitable set of matings, in order to achieve his/her breeding objectives taking into account the restrictions on budget imposed.

A cost-benefit analysis of implementation has not been attempted in this case.

There seems to be little or no previous work similar to the analysis just carried out, probably because of a relative lack of development of mate selection systems that operate beyond control of genetic merit and coancestry.

#### **4.4.2 Case B) Dissemination. Dispersal of breeding males to different beef production units oriented to diverse beef markets**

Table 4.4 shows the final results of this case.

Table 4.4 Results for the different analysed objectives  
(see section 4.3.3.2.3 to refresh the meaning of the terms used in this table)

Breeding Objective		a		b		c	
		Balanced target (25 degrees)		Minimum Coancestry (90 degrees)		Minimum of F (0 degrees)	
Item	Market	Average Index	Export / Domestic	Average Index	Export / Domestic	Average Index	Export / Domestic
Achieved Mean Progeny Index (\$)		39.969	40.023	26.171	26.267	40.170	40.174
Achieved Coancestry		0.0736		0.0600		0.1248	
Achieved Progeny Mean F		0.0003		0.0003		0.0000	

As observed in Table 4.4, a reduction in the levels of coancestry corresponded to a reduction in genetic gain in the next generation in both strategies. There were not significant differences in short-term inbreeding between scenarios. Thus, it can be observed that generating maximum genetic diversity caused a reduction in genetic merit (profit) of around 52% in average in both strategies compared to use of a balanced strategy of 25 target degrees (breeding objective ‘a’ vs. ‘b’).

In the three scenarios of the strategy in which an average of the indices was used in both herds, genetic merit for profit was smaller than when these indices were used separately in the corresponding herd. Even though the differences between the achieved genetic gains of the analysed strategies were small, the use of separated indices in each herd appeared as a more effective approach. It should be noted that these differences will be smaller as the correlation between the Export market and Domestic market indices increases. In the present case, the correlation between these indices was 0.925. This agrees with the trend shown in Figure 4.8, which shows the difference between the resulting average genetic merit of progeny of the two strategies considered as a function of the correlation between the Export market and Domestic market indices. Different correlations were generated by random sampling of the

indices under the two markets. It should be noted that the resulting differences shown also relate to the size and structure of the database used in this section and the sampling errors will be involved.

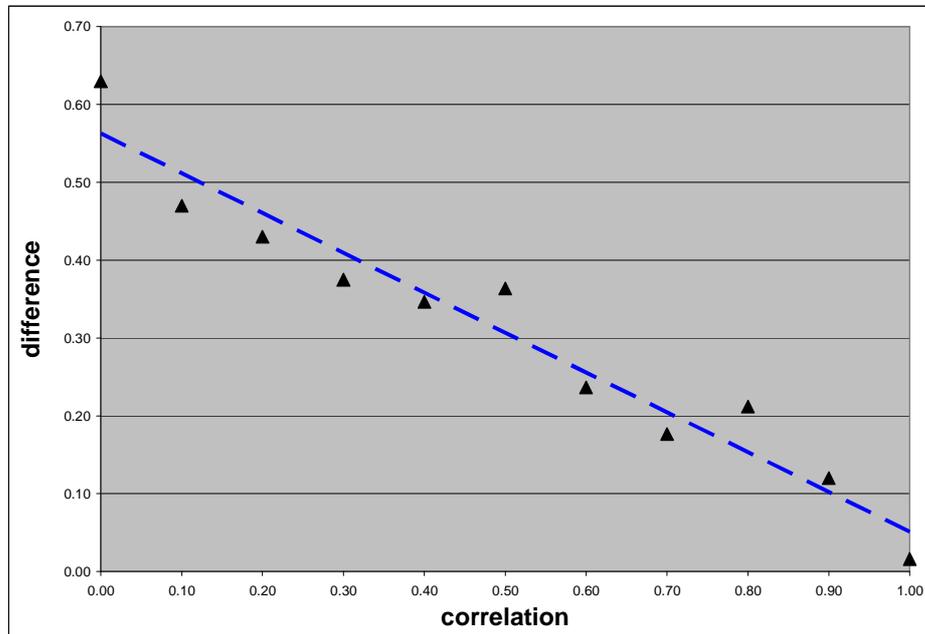


Figure 4.8 Differences between the resulting average genetic merit of progeny of two Mate Selection analyses (setting 25 target degrees), one obtained when the Export market and Domestic market indices were used separately and the other when the average index was used, according the correlation between the Export market and Domestic market indices in the analyses (trend = slashed line).

In order to achieve the optimum genetic merit those sires with the highest indices were selected, considering the chosen constraints in each scenario.

On the other hand, when the Export market and Domestic market indices were used separately, optimisation of genetic merit in each economic index was achieved in different herds of cows, with appropriate competition between cow herds for sires that are high for both indices. In this way, it was shown that dispersal of sires to different breeding units can be handled simultaneously with selection of sires into the stud, even considering additional constraints.

#### 4.4.3 Case C) Mate Selection of breeding males and females to obtain a bimodal population oriented to two beef markets.

In the following Table 4.5 are shown the final results of the analyses of this case. In the first part of this table the results for the different strategies are presented as in the previous case B but in the second part below they were divided as:

- Avg Ind (E): those that correspond to when the average index was used and the mentioned replacement in the previous ‘Analysis method’ section was done with the corresponding Export market index of candidates;
- Avg Ind (D): as above but for the Domestic market index;
- E: those that correspond to when the Export market index was used separately in the strategy;
- D: as above, but for the Domestic market index.

Table 4.5 Results for the different analysed objectives

Breeding Objective		a		b		c	
Item	Market	Average Index	Export / Domestic	Average Index	Export / Domestic	Average Index	Export / Domestic
	Proportion of herd		0.2 / 0.8		0.5 / 0.5		0.8 / 0.2
Achieved Mean Progeny Index (\$)		40.93	41.03	42.27	42.45	43.72	43.80

Sire ID	Uses				Uses				Uses			
	Avg Ind		E	D	Avg Ind		E	D	Avg Ind		E	D
	(E)	(D)			(E)	(D)			(E)	(D)		
6985	2	5	–	32	3	4	–	27	7	–	–	6
8554	22	88	30	81	55	55	75	37	107	3	110	–
10145	6	28	–	8	17	17	–	12	6	28	10	25

As observed in Table 4.5, considering the three ratio scenarios ('a', 'b' and 'c'), it can be observed that there was a small but progressive increment in mean progeny genetic merit from 'a' to 'c' for both strategies.

In order to achieve the optimum genetic merit those sires with the highest indices were selected, considering the chosen constraints in each scenario.

Although the same sires were used in the different resulting sets of matings (see lower part of Table 4.5), the proportion and distribution of the use of each of them varied along the scenarios and strategies. Thus, the progressive increment of the mean progeny index just mentioned in the previous paragraph was caused by the progressively higher use of the candidate sire ID 8554, which was the sire with the highest export index, mated to the females of higher export index.

In the three scenarios of the strategy in which an average of the indices was used, genetic merit (profit) was lower than when these indices were used separately in the herd. Even though the differences between the achieved genetic gains of the analysed strategies were small, the use of separate indices in the herd was more efficient. It should be noted that these differences will be smaller as the correlation between the Export market and Domestic market indices increases (see Figure 4.8).

In this way, when the indices were used separately, optimisation of each economic index was achieved in a herd of cows varying the proportions that were focused to different indices, i.e. Export market and Domestic market indices. Thus, it was shown that dispersal of sires to different breeding units can be handled simultaneously with selection of sires into the stud, even considering additional constraints.

#### **4.4.3.1 Conclusions**

The results obtained in the last two cases presented showed that the application of Mate Selection is a tactical approach suitable to simultaneously

select different sets of sires in the stud-herds according the particular breeding objectives of different commercial units, even including those of the stud. The consideration of these objectives in the MSI function allows the optimisation of profit over herds, in addition to the control of other genetic aspects such as maintenance of genetic diversity and avoidance of the undesirable consequences of inbreeding, and, as shown in the first case, appropriate exploitation of reproductive technologies.

Although the results must be considered illustrative rather than evaluative, the link between mate selection method and results achieved is direct, and less subject to sampling error because contrasted treatments use the same dataset.

There seems to be little or no previous work similar to the analyses just carried out, probably because of a relative lack of development of mate selection systems that operate beyond control of genetic merit and coancestry.

## **CHAPTER 5**

# **DETERMINISTIC OPTIMISATION OF A YOUNG SIRE PROGRAM**

## ***5.1 Summary***

In this chapter deterministic modelling is used to explore the balance between selection of old versus young sires as nucleus parents in a YSP. The mean and variance of all age cohorts of males and females are simulated before and after selection, and optimised selection across cohorts based on the objective of either maximising next generation response (no-LAMS) or the selection response that can be achieved two generations ahead through a Look-Ahead Mate Selection approach (LAMS). How this balance is affected by various parameters related to breeding program design is explored, in particular the differences in accuracy of EBV in various age classes.

The use of a DE algorithm was the key tool to determine the optimal proportions of proven and young bulls to be selected as parents along age classes and balance genetic parameters.

A higher increase in the number and proportion of selected animals in the age class with higher accuracy occurs in no-LAMS approach, while in LAMS approach it is higher in the age class with lower accuracy. The LAMS approach results in a slightly higher long-term genetic response compared to aiming to improve the immediate next generation (no-LAMS). This marginally higher rate of genetic gain increases more as more young bulls are progeny tested, increasing in this way the number of sire candidates across all age classes. Around 3.0% more genetic gain was obtained in LAMS compared to no-LAMS when 100 proven bulls are available as candidates to be selected as sires in both approaches. In addition, it is observed that the age distribution of sires is also affected, with more proportion of young sires selected as nucleus sires under LAMS than in no-LAMS.

## ***5.2 Introduction***

A considerable proportion of stud-farms with a limited herd size exists in Argentina. For each stud individually it is not feasible to run a separate breeding program and some form of collaboration or exchange of genetic material is required. An efficient way to work together is a Young Sire Program (YSP). YSP allows breeders to share risk and opportunity by testing progeny of young bulls in each others herd. This provides an opportunity to progeny test more young bulls, which would be an expensive exercise for individual studs, and this could improve the rate of genetic gain with less increase in inbreeding. The program would also link up herds very well, which is an important aspect of across herd sire evaluation. In addition, commercial clients benefit from more accurate breeding values and better genetics and a wider use of local leading young bulls decreases the cost of genetic progress compared with the use of semen of outside proven bulls.

A Young Sire Program involves mating young sires to generate test progeny. Comparing young sires with older proven sires, the former are usually cheaper as less information is known about them and their estimated breeding values are less accurate. On the other hand, the older sires have estimated breeding values with more accuracy as their progeny have records, but older sires are born in a cohort that should have a lower genetic mean than the more recent cohorts of young sires. Using young males reduces generation interval, which in turn conduces to the acceleration of the rate of genetic gain.

A dilemma in YSP is the proportion of matings that should be reserved for young sires. Young sire matings could take place outside the nucleus, i.e. to the inferior cows in each stud. The number of matings outside the nucleus could be increased in order to progeny test more young sires, creating in this way the progeny testing capacity. Extra progeny testing capacity outside the nucleus is needed to test good young bulls and to avoid having only a very limited number

of proven bull candidates to compete with the new young bulls to be mated in the nucleus.

Some authors have examined the optimal proportions of young and old sires. Banks et al. (1998) carried out a simple deterministic simulation in sheep breeding to examine the optimal proportion of young to old sires with different amount of information which results in differences in accuracy. In that study optimisation was done through selection on BLUP index across age groups, including only 2 age groups of males and a fixed age structure of females. The effects of varying heritability and size of the population were also examined. On the other hand, Meszaros et al. (1999) have used genetic algorithms to optimise age structure of females and mating ratio in sheep breeding programs. That work considered the effect of fixed inbreeding levels on depression of genetic gain through effects on reproduction and variation of the population size.

In this study the ratio of young to proven sires to be used in nucleus matings will be optimised but a varying number of young bulls progeny tested outside the nucleus. For nucleus matings, the application of truncation selection based on BLUP EBV would optimise the selection of sires in the sense that it gives the highest genetic mean of their progeny (James, 1987). The same author showed that this type of selection optimises generation interval, providing an optimal allocation of selected sires across the eligible age classes. This involves simply selecting the sires with the highest breeding values. However, if relatively more young sires were selected, the variation among sires' true breeding values would increase as young sires are selected with less accurate estimated breeding value but with more prediction error variance (PEV). This could give more genetic gain over the longer term, as suggested by Kinghorn and Shepherd (1994). Note that the distribution of true breeding value given an estimated breeding value (EBV) is normal with a mean equal to EBV and a variance equal to PEV (see Figures 5.1 and 5.2).

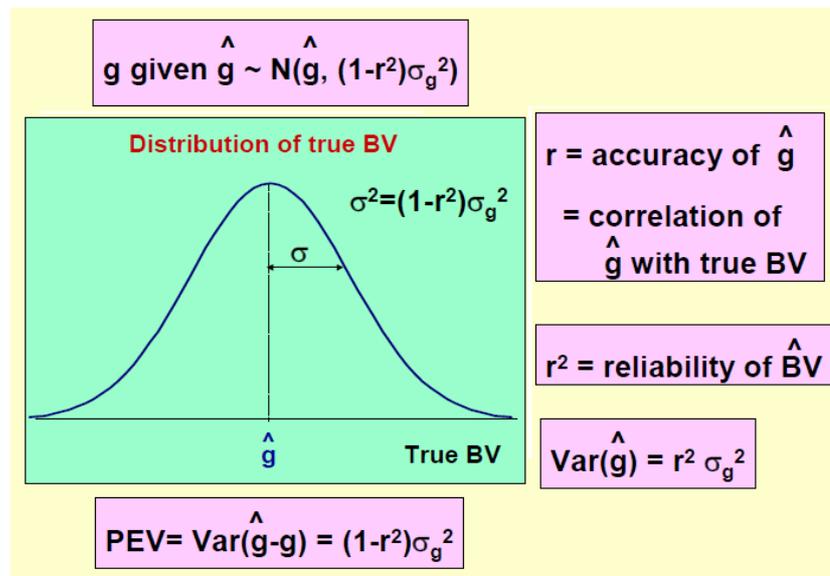


Figure 5.1 Distribution of True BV given EBV, where  $g$  refers to True BV,  $\hat{g}$  to EBV,  $r$  to the accuracy of EBV, PEV to the prediction error variance,  $\sigma_g^2$  to the genetic variance. (Gibson and Dekkers, 2009)

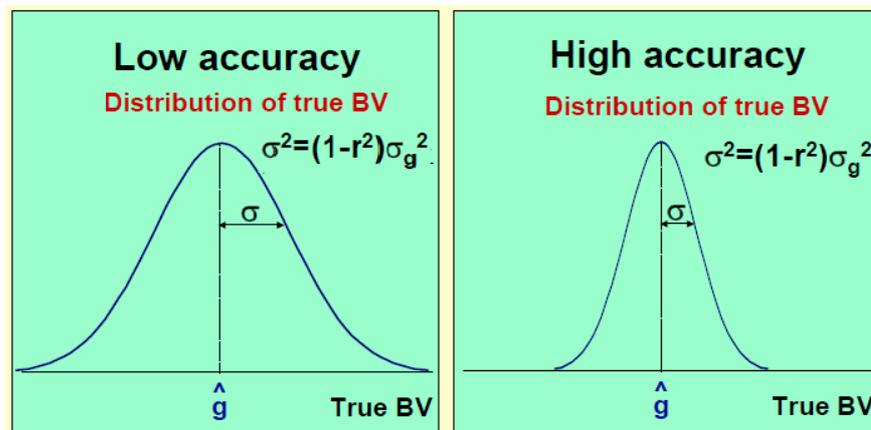


Figure 5.2 Effect of accuracy on distribution of True BV for animals with a given EBV, where  $\hat{g}$  refers to EBV,  $r$  to the accuracy of EBV,  $\sigma_g^2$  to the genetic variance and  $\sigma^2$  to the prediction error variance (PEV). (Gibson and Dekkers, 2009)

More variation would result in more selection differential and if the mean is not compromised too much this could lead to more genetic gain in the long term.

Based on this proposition, Kinghorn and Shepherd (1994) suggested the use of Look-Ahead Mate Selection (LAMS) in which the aim of the selection

process is to maximize the predicted genetic selection differential in the next generation rather than in the current generation. Or, in other words, as it is described in detail later in section 5.4.1.3, the objective is to maximize, with the current selection round, the mean expected breeding value of the grand progeny rather than of the progeny. The selection differential in the progeny generation can be predicted from the difference between the predicted average of the animals selected to become parents and the average of the entire population of prospective future progeny. This selection differential depends on the variance of the true BV, which is constituted from the prediction error variance and the variance among cohorts, assuming parents are selected each year from multiple cohorts (age classes). Note that the evaluation of individuals is not delayed by a generation, but a prediction of second generation impact is immediately available such there is no deleterious impact via longer generation intervals.

In this chapter deterministic modelling will be used to explore the balance between selection of old versus young sires as nucleus parents in a YSP. The mean and variance of all age cohorts of males and females are simulated before and after selection, and optimised selection across cohorts based on the objective of either maximising next generation response (no-LAMS) or the selection response that can be achieved two generations ahead. How this balance is affected by various parameters related to breeding program design is explored, in particular the differences in accuracy of EBV in various age classes and the number of young bulls that gets progeny tested.

### ***5.3 Objectives of the analysis***

The main objective of the study is to find out if Look-Ahead Mate Selection (LAMS) could give more genetic gain on the longer term than truncation selection and to check whether in an optimal breeding program the proportion of young sires selected should be increased compared to common

practices. First, the approach will be described in general terms, while a more detailed description will be given in the following sections.

Sets of genetic parameters will be varied, e.g. selection accuracies of different age classes, and these will be used in a deterministic model of a breeding nucleus in order to measure and understand the impact of those parameters on the achieved outcomes and the number of young bulls that gets progeny tested. Optimal solutions for various scenarios in terms of proportion of young bulls selected for the nucleus and number of young bulls mated for the purpose of progeny testing are shown in this analysis. In addition, the differences in rate of genetic improvement are compared between traditional breeding systems in which sires are selected by truncation selection, i.e. looking only one generation ahead versus alternative breeding schemes which aim to maximize the predicted genetic selection differential in the next generation, i.e. looking two generations ahead.

#### ***5.4 Materials and methods***

A deterministic model was developed to predict the response to selection with overlapping generations. In the first year,  $N$  females were distributed equally over 5 age classes. The number of males was equal to  $N/F$  where  $F$  is the number of females mated to each male. Initially the mean breeding value (BV) in each age class was zero. The variance of BV ( $\sigma_g^2$ ) was 1 and the variance of the EBV of each age class was equal to  $r^2\sigma_g^2$  according to the corresponding accuracy ( $r$ ) of each age class. In the subsequent years, the selection differential was calculated as the sum of the products of  $p$  by the genetic mean of sires and dams selected in each age class, where  $p$  refers to the proportion of selected parents originated from that age class while the mean is the genetic mean of the selected proportion of each cohort which contributes to the total number selected parents. The optimal selection differential was

calculated by optimising the proportion selected from each age class. For no-LAMS, this optimisation was based on truncation selection across age classes of males and females, respectively.

Optimised selection across age classes can be achieved by using the algorithm of Ducrocq and Quaas (1988). This algorithm determines the threshold across age classes ( $TH$ ) and from that the proportion selected to each age class can be calculated, the selection intensity ( $i$ ) and selection threshold ( $x$ ) which is standardized within the distribution of each age class as  $(TH - \mu_j) / \sigma_j$ , where  $\mu_j$  is the predicted genetic mean of the selected group of males and females from age class  $j$  and  $\sigma_j$  is the standard deviation of estimated breeding values within the age class  $j$ . Such truncation selection can be used to maximize response to the next generation. Note, however, that this approach assumes normal shape of distributions for derivation of variances, intensities and thresholds in all but the first round. The fact that this assumption ignores possible deviation from normality after selection helps motivate the stochastic analyses carried out in the next Chapter.

In deterministic LAMS two generations ahead are looked by maximizing the predicted genetic selection differential in the next generation.

In this chapter, the use of an evolutionary algorithm was proposed to find optimal solution for the proportions selected from each age class. As it was already mentioned, simulations were done over multiple selection rounds, being the approach fully tactical as the solution was optimised at every selection round. A deterministic model was set in an Excel spreadsheet which used a macro written in Visual Basic language. The model was able to predict response to selection in a seed-stock beef herd for a given strategy, determined by the proportion selected from each age class. Based on the proportion selected the mean and variance of the selected animals in that age class could be calculated. In turn, based on the proportional contribution of each age class to the progeny in the next generation, the mean and variance of the BV of the progeny could be calculated, which became the first age class in the next year. The initial accuracy

of EBV in age class 1 depends on the information available at that age. These accuracies were varied for different scenarios as detailed in the next sections. Accuracies in later age classes are affected by a reduction in genetic variance due to selection (the ‘Bulmer effect’; Bulmer, 1971).

The model uses a Differential Evolution algorithm (Storn and Price, 1997) to find the optimal proportion selected according to a certain criterion (objective function). A DE algorithm is an efficient search algorithm for complex problems. In this case, there are many possible solutions from the very many combinations that can be made from proportion selected in each age class. A possible solution is a set of values that reflect the proportions selected from each age class in males and females, being this solution optimised at every selection round. In the DE algorithm, proposed solutions need to be evaluated by an objective function, whose corresponding details are given further in the next sections. The objective function will relate to the achieved rate of genetic gain, either by maximizing the predicted mean of progeny, or that of grand progeny.

It is important to mention that, even though in two of the next presented cases simply truncation selection for optimisation could be used, as the use of a DE algorithm is not strictly needed, its use is needed in the third scenario. Therefore, the DE algorithm was used for all cases to keep consistency across the analysis.

#### **5.4.1 Scenarios**

The different scenarios with increasing complexity to be analysed in this chapter were:

Case A) Optimal proportions of selected sires in a herd with two age classes for males assuming differences in EBV accuracy between these age classes.

Case B) Optimal selected proportions of sires with respect to mean of progeny in a herd with several age classes for males and females with varying initial accuracy for the 2-year-old age class of sires and the test capacity.

Case C) As case B) but now using the LAMS approach, i.e. optimising the mean of grand progeny.

#### **5.4.1.1 Case A) Optimal proportions of selected sires in a herd with two age classes for males assuming differences in EBV accuracy between these age classes.**

We started with a model for a simple problem considering two age classes, i.e. young bulls vs. proven bulls, and optimising their proportions for a given difference in accuracy and age. Basically there is only one parameter to optimise (as number of proven bulls selected implies number of young bulls for a given total).

The genetic response ( $R$ ) is calculated based on Rendel and Robertson's formula.

The response is calculated based on the average selection differential of parents divided by the average generation interval, where the selection differential achieved within a cohort is calculated as:

$$S_j = i_j * r_j * \sigma_g \quad (5.1)$$

where  $j$  refers to the cohort (e.g. age class from one sex),  $i$  refers to the respective selection intensity,  $r$  is the corresponding accuracy and  $\sigma_g$  is the genetic SD.

The response to selection where a mixture of young and proven bulls is used is then:

$$R = \frac{p_{YB}i_{YB}r_{YB} + p_{PB}i_{PB}r_{PB} + i_f r_f}{p_{YB}L_{YB} + p_{PB}L_{PB} + L_f} * \sigma_g \quad (5.2)$$

where  $p_{YB}$  and  $p_{PB}$  refer to the proportion (out of total bulls selected) of young and proven bulls, respectively. The parameters  $i_{YB}$  and  $i_{PB}$  correspond to the selection intensity,  $r_{YB}$  and  $r_{PB}$  to the accuracy and  $L_{YB}$  and  $L_{PB}$  to the generation interval of young and proven bulls, respectively while  $\sigma_g$  is the additive genetic standard deviation of the population. The subscript  $f$  refers to females (dams).

Initially, only the parameter  $p_{YB}$  needs to be optimised, as  $p_{YB} + p_{PB}$  add up to 1. For this, the proportion of proven bulls selected out of all proven bulls (called  $prop_{PB}$ ) is optimised whereas other parameters, such as  $prop_{YB}$ ,  $i_{YB}$ ,  $i_{PB}$ ,  $p_{YB}$  and  $p_{PB}$ , are derived from that. Although a simple linear search is possible for this problem, a differential evolution algorithm following Price and Storn (1997) was used which optimises only 1 parameter by trialling values for  $prop_{PB}$  and evaluates each value by an objective function, which is simply the calculated value for  $R$ .

EBVs were assumed to be based on BLUP, which implies that they are comparable across age classes.

Figure 5.3 shows the effect of accuracy on the distribution of estimated breeding values. This is essential when selecting across age classes, where the accuracy (and therefore the SD of the estimated breeding value) varies between cohorts.

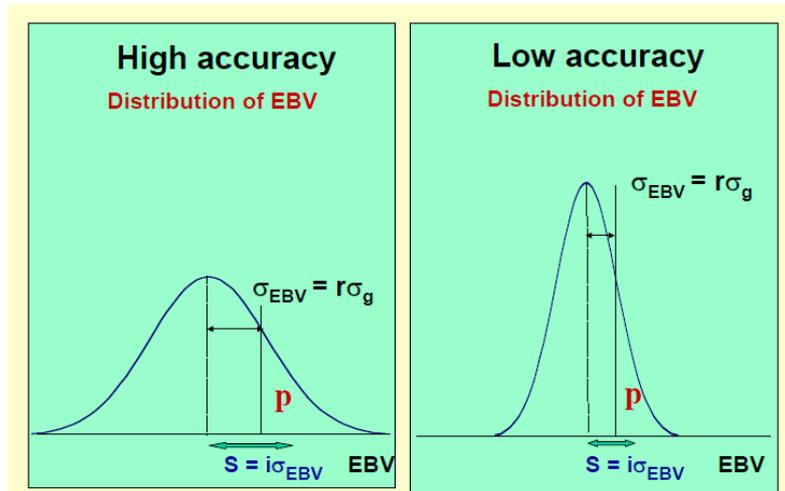


Figure 5.3 Effect of accuracy on distribution of EBV, where  $p$  refers to the selected proportion of parents and  $S$  to the selection differential. (Gibson and Dekkers, 2009)

The example shown in Figure 5.4 is modelled for a closed breeding nucleus herd with in total 1000 females that supplies bulls to a multiplier/commercial tier, producing 500 young bulls. Annually, 25 of those 500 young bulls are progeny tested outside the nucleus in order to have 25 proven bulls available to select from as nucleus sires and to compete with the 500 young bull candidates. These proportions are equivalent to using a mating ratio equal to 40 dams per sire. Those young bulls that are not selected for nucleus use and not progeny tested will be culled.

In the model, some assumptions are made about genetic parameter differences in the herd, e.g. genetic mean and variance. An essential parameter to be considered is the difference in mean BV between young bulls and proven bulls. This difference is equal to the age difference (i.e. 2 years) times the genetic progress per year ( $R$ ). Hence, an iterative process is needed to find the optimum selection differential to maximize  $R$ . Note that both age classes were selected on EBV, with the EBV variance calculated as  $r^2\sigma_g^2$  (i.e. the squared accuracy (which is also referred as reliability) times the genetic variance). It was not considered to give proven bulls a reduced genetic variance due to selection (although in practice they would have lower genetic variance due to the Bulmer



To make the analysis more realistic for a YSP, the developed spreadsheet was extended to more age classes (5 in the example shown in Figures 5.5a and 5.5b) not only for males but also for females. Now, the closed herd includes a breeding nucleus with a total of 1000 combined females from 10 small studs participating in this YSP. These females are used for progeny testing of young bulls (age classes 2 to 4) and proven bulls (age class 5), which in turn are used as selection candidates. It is important to mention that young bulls are also tested outside the nucleus. A mating ratio equal to 40 dams per sire was used. A 5% mortality was assumed, which affects the total available number of males and females in each age class (i.e. selection candidates). The minimum age at mating for males and females was set to 2 years old. The breeding program aims to produce 25 progeny tested bulls. Since progeny are not tested until year 5, and because of the 5% mortality rate, 33 young bulls are mated in year 2 for progeny testing, leaving 30, 27 and 25 bulls in age classes 3, 4 and 5. Young bulls not mated for progeny testing were culled after year 2. Young bulls are selected to be progeny tested based on parent average genetic merit or also on own performance. It is also necessary to mention that, in cases B and C, it was considered one selection round per year (i.e. a mating season), and one year difference between age classes.

The Rendel and Robertson formula to predict response in the next generation can still be used, by expanding / generalizing Eqn. [5.2]:

$$R = \sum_{j=1}^{nr \text{ of cohorts}} \frac{p_j i_j r_j}{\sum_{j=1}^{nr \text{ of cohorts}} p_j L_j} * \sigma_g \quad (5.3)$$

where  $p$  is the proportion of each cohort contributing to the total number selected (note that these values should add up to 1 for each sex).

As pointed out by James (1987), an optimal solution for  $R$  in Eqn. [5.3] is found by truncation selection across age classes, provided that the selection

criteria are comparable across age classes. This is the case with BLUP breeding values as they account for genetic trend. Rather than optimising the average of selection differentials within an age class along with the average generation interval, the average of the mean breeding value of all selected parents can be optimised. Note that the mean breeding value of selected parents from an age class is the sum of selection differential within the age class ( $S_i$ ) and the annual genetic trend difference between age classes. If the mean of selected parents from the oldest age class is equal to  $S_i$ , then the mean of a younger age class is  $S_i + a \cdot dG$ , where  $a$  is the difference in years between the age classes and  $dG$  is the genetic trend per year. Hence, it can be shown that instead of maximizing Eqn. [5.3], the genetic mean of selected parents can be maximized, as in Eqn. [5.4] below.

The DE algorithm used in the macro was modified to maximize the value for the genetic mean of selected parents ( $S$  in Eqn. [5.4]) by trialling values for  $pm$  of males and for  $pf$  of females along all age classes (i.e. the number of males (females) selected out of each age class as a proportion of the total number of males (females) selected and evaluating these values by an objective function, which is the average expected breeding value of the selected males and females as parents of the next generation ( $S$ ):

$$S = \left( \sum_{j=1}^{nr \text{ of } malecohorts} pm_j * meanm_j + \sum_{k=1}^{nr \text{ of } femalecohorts} pf_k * meanf_k \right) / 2 \quad (5.4)$$

where  $S$  is the selection differential of the selected males and females as parents of the next generation, and  $meanm_j$  and  $meanf_k$  is the genetic mean of the proportion of each cohort of males ( $pm_j$ ) and females ( $pf_k$ ) respectively contributing to the total number selected (see Figure 5.6).

PREDICT OFFSPRING FOR A GIVEN SELECTION					input variables in light blue cells					Run evolutionary algorithm				
nr years	50	nreps	1											
nr age classes	min age	SDa gen 0	Mating ratio	Females										
5	2	1	40	1000										

1st mat season	number	selected	mean	P	varG									
Age classes	in nucleus	number	varG	mean	prop	i	accuracy	Self if	selected	prop of selected	selected	PEV	SEP	
Young Bulls 1	500		1.00	0.00			0.40							
Young Bulls 2	475	17.72	1.00	0.00	0.037	2.18	0.46	1.00	1.00	0.709	0.82	0.79	0.89	
Young Bulls 3	30	1.57	1.00	0.00	0.052	2.04	0.51	1.04	1.04	0.063	0.78	0.74	0.86	
Young Bulls 4	27	1.90	1.00	0.00	0.070	1.92	0.56	1.07	1.07	0.076	0.73	0.69	0.83	
Proven Bulls 5	25	3.82	1.00	0.00	0.153	1.55	0.80	1.24	1.24	0.153	0.49	0.36	0.60	
Females 1	500		1.00	0.00			0.40							
Females 2	475	271.8	1.00	0.00	0.572	0.69	0.42	0.29	0.29	0.272	0.90	0.82	0.91	
Females 3	451	256.5	1.00	0.00	0.569	0.69	0.44	0.30	0.30	0.257	0.88	0.81	0.90	
Females 4	429	242.6	1.00	0.00	0.566	0.70	0.46	0.32	0.32	0.243	0.87	0.79	0.89	
Females 5	407	229.0	1.00	0.00	0.563	0.70	0.48	0.34	0.34	0.229	0.86	0.77	0.88	
Response	0.679													

Figure 5.5a Overview of selection activity in each age class for case ‘B’ (initial mating season), where ‘*varG selected*’ is the genetic variance of the selected parent of each cohort, *prop* is the proportion of parents selected from that age class, *p* is the proportion of parents from each age class as proportion of total number of parents used and *SEP* is the standard error of prediction (i.e. the squared root of *PEV*) of each cohort.

20th mat. season	number	selected	mean	P	varG									
	in nucleus	number	varG	mean	prop	i	accuracy	Self if	selected	prop of selected	selected	PEV	SEP	
Young Bulls 1	500		0.91	4.14				0.00	0.00					
Young Bulls 2	475	22.22	0.91	3.95	0.047	2.09	0.37	0.76	4.71	0.889	0.81	0.79	0.89	
Young Bulls 3	30	0.83	0.91	3.75	0.028	2.30	0.43	1.00	4.74	0.033	0.76	0.74	0.86	
Young Bulls 4	27	0.50	0.91	3.56	0.018	2.45	0.50	1.21	4.77	0.020	0.71	0.69	0.83	
Proven Bulls 5	25	1.45	0.91	3.35	0.058	2.00	0.78	1.56	4.91	0.058	0.44	0.36	0.60	
Females 1	500		0.91	4.14					0.00					
Females 2	475	411.7	0.91	3.95	0.867	0.25	0.31	0.08	4.02	0.412	0.88	0.82	0.91	
Females 3	451	296.3	0.91	3.75	0.657	0.56	0.34	0.19	3.94	0.296	0.85	0.81	0.90	
Females 4	429	187.5	0.91	3.56	0.437	0.90	0.37	0.33	3.89	0.188	0.83	0.79	0.89	
Females 5	407	104.5	0.91	3.35	0.257	1.26	0.39	0.49	3.85	0.104	0.80	0.77	0.88	
R 20th mat. season	0.197													

Figure 5.5b Overview of selection activity in each age class for case ‘B’ (after 20 years (rounds) of selection, where equilibrium has already been reached). Compared to Figure 5.5a, observe the lower genetic variance of each cohort, higher than zero mean of each cohort and lower genetic response.

Sex				Bulls			
	age class			2	3	4	5
		prop of selec	mean of selec	pm(2)	pm(3)	pm(4)	pm(5)
				0.889	0.033	0.020	0.058
				meanm(2)	meanm(3)	meanm(4)	meanm(5)
				4.713	4.744	4.771	4.908
Females	2	pf(2) 0.412	meanf(2) 4.025	1.599	0.060	0.036	0.107
	3	pf(3) 0.296	meanf(3) 3.938	1.139	0.043	0.026	0.076
	4	pf(4) 0.188	meanf(4) 3.886	0.717	0.027	0.016	0.048
	5	pf(5) 0.104	meanf(5) 3.847	0.398	0.015	0.009	0.027
selection differential of selected parents							4.340

Figure 5.6 Example of the calculation of  $S$  using the results shown in Figure 5.5b and the formula  $(\sum pm_j * meanm_j + \sum pf_k * meanf_k) / 2$ .

This value ( $S$ ) in Eqn. [5.4] is the predicted mean of age class 1 of the next generation.

It can be observed in Figure 5.5a that different parameters can be used as initial values (those highlighted in light blue cells) in order to obtain results. The initial values for accuracy depend on the measurement strategies used; more traits measured at a young age would give higher accuracies for young bulls. Also, when genomic selection is used, it would increase accuracies of young bulls relatively more. Low and high values for initial accuracy for the 2-year-old age class of sires were tested to study their impact on the optimal selection scheme. Taking into account the decision of carrying out a YSP, these changes were mainly directed to know whether and how many young sires should be progeny tested.

The mating ratio used implies that a total of 25 males and 1000 females need to be selected each year as sires and dams in this nucleus composed by a group of small herds. The total numbers of young (age classes 2 to 4) and proven (age class 5) candidates for males were 532 and 25 respectively. While a

total of 1355 young (age classes 2 to 4) and 407 older (age class 5) candidates for females were used.

Initially, all age classes have the same mean but, due to selection and genetic trend, differences will appear. An equilibrium in the age structure of genetic gain was reached by running a selection program for 50 years, each time calculating the mean and variance of selected parents within each cohort for males and females along all age classes.

It is necessary to note that, in this case B (no-LAMS), in order to be consistent to the following case C (LAMS), the annual rate of predicted genetic gain is calculated to make comparisons as the difference between the predicted mean of age class 1 and age class 2 each year.

As already mentioned in section 2.1.4 of Chapter 2, it is important to consider that selection has an impact not only on the mean of the population but also on genetic variance. Thus, genetic variance is reduced by selection due to the Bulmer effect. Changes in genetic variance affect the amount of change that can be made in future generations.

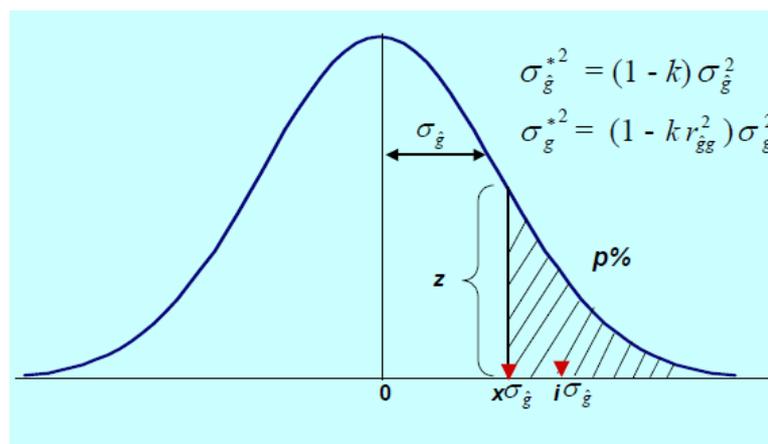


Figure 5.7 Selection on EBV ( $\hat{g}$ ), where  $\sigma_{\hat{g}}^2$  is the variance in EBV,  $\sigma_g^2$  is the genetic variance, \* is used to denote parameters after selection,  $i$  is the selection intensity,  $x$  is the truncation point,  $k$  is the variance reduction coefficient calculated as  $i \cdot (i - x)$ ,  $r_{\hat{g}g}$  is the correlation between EBV ( $\hat{g}$ ) and TBV ( $g$ ) and  $p$  is the proportion of selected animals. (Gibson and Dekkers, 2009)

The impact of the Bulmer effect was accounted for in the code of the program used for this analysis. The reduction in variance among selected individuals can be derived from the genetic theory (Bulmer, 1971). The corresponding formulae can be seen in Figure 5.7.

The genetic variance of a population prior to selection can be partitioned into the parental and Mendelian sampling components as:

$$\sigma_g^2 = 0.25\sigma_{gs}^2 + 0.25\sigma_{gd}^2 + \sigma_{gm}^2 \quad (5.5)$$

where  $\sigma_g^2$  is the genetic variance of a population prior to selection,  $\sigma_{gs}^2$  and  $\sigma_{gd}^2$  is the genetic variance of sires and dams respectively and  $\sigma_{gm}^2$  is the variance due to the Mendelian sampling.

The formulae shown above can be modified to give genetic variance of animals after selection of sires and dams and generalized to predict the genetic variance in generation  $t+1$  from the variance among the parents selected in generation  $t$ :

$$\sigma_{g(t+1)}^2 = 0.25\sigma_{gs(t)}^{*2} + 0.25\sigma_{gd(t)}^{*2} + \sigma_{gm}^2 \quad (5.6)$$

where  $\sigma_{g(t+1)}^2$  is the genetic variance of a population in generation  $t+1$ ,  $\sigma_{gs(t)}^{*2}$  is the genetic variance among the selected sires in generation  $t$  and  $\sigma_{gd(t)}^{*2}$  is the genetic variance among the selected dams in generation  $t$ .

It is necessary to note that only the parental contributions to variation are affected by selection. The variance due to the Mendelian sampling,  $\sigma_{gm}^2$ , is unaffected by selection and is equal to  $\frac{1}{2}\sigma_{g(0)}^2$ , where  $\sigma_{g(0)}^2$  is the genetic variance in the unselected and non-inbred base population. The Mendelian sampling variance would reduce due to inbreeding, but inbreeding will be not considered in the analysis of this chapter.

Based on the previous formulas, the following general equation can be developed to predict genetic variance among progeny:

$$\sigma_{g(t+1)}^2 = 0.25(1 - k_s r_{s(t)})\sigma_{gs(t)}^{*2} + 0.25(1 - k_d r_{d(t)})\sigma_{gd(t)}^{*2} + 0.5\sigma_{g(0)}^2 \quad (5.7)$$

where  $k_s$  and  $k_d$  are the variance reduction coefficient of sires and dams respectively, with being  $k$  calculated as  $(i - (i - x))$  where  $i$  is the selection intensity and  $x$  is the truncation point used for selection in an age class, and  $r_{s(t)}$  and  $r_{d(t)}$  are the respective accuracies of selection in sires and dams in generation  $t$  (Gibson and Dekkers, 2009). The reduction in variance in the new cohort of progeny born translates also in a change in accuracy.

In case B, a base scenario is first presented with low initial accuracy (due to e.g. few traits measured) for the 2-year-old age class and the optimal proportion of bulls from each age class is observed. Then, another scenario is presented with a higher initial accuracy for the 2-year-old age class, e.g. due to using genomic information or measuring more traits early on. It is necessary to note that the procedures that allow increase or improve the information on new candidates (e.g. DNA analyses in male calves for genomic information or measurement of more traits in the 2-year-old age class) are done every selection round. This determines an increase in the initial accuracy for the 2-year-old age class and, in turn, also of the accuracies for the following age classes (i.e. from 3 to 5-year-old age classes). Thus, the resulting optimal proportions between these scenarios are compared. A different set of accuracies was used for females in all scenarios (see Table 5.1). Note also in Table 5.1 that a relatively higher difference in accuracy exists between the last sire age class (5-year-old age class) and the younger sire age classes (2 to 4-year-old age classes) because sires of 5-year-old age class have been progeny tested (proven sires).

Table 5.1 Sets of initial accuracies used for bulls and females

Age classes	initial accuracies for bulls		initial accuracies for females
	measuring few traits	measuring more traits or genomic information	
1	0.40	0.46	0.40
2	0.46	0.52	0.42
3	0.51	0.57	0.44
4	0.56	0.62	0.46
5	0.80	0.81	0.48

Then, the consequences of varying the test capacity were analysed (i.e. the number of young bulls (age classes 2 to 4) that are progeny tested annually or progeny per young bull tested per mating season). For this, first the results of having 25 proven bulls at year 5 were analysed, and two alternative numbers of proven bulls as candidates were examined, i.e. 100 and 200. This is achieved by mating annually these numbers of 2-year-old young bulls to cows outside the nucleus. This means that annually the candidates of 3 and 4-year-old age class bulls also increase to those numbers. Thus, when these alternative numbers of proven bull candidates are available, the total of young (age classes 2 to 4) candidates for males are 690 and 905 respectively. As this extra number of cows for progeny testing does not belong to the nucleus they are not modelled, so the total young and older female candidates remains the same. It should be noted that the mating ratio used is not changed either, i.e. it is of 40 dams per sire. Finally, the results of the base scenario were compared to those using a different mating ratio, i.e. 30 dams per sire. This is an alternative way to create more progeny tested bulls, but with fewer progeny per bull and without having to increase the progeny test capacity outside the nucleus. It should be noted that although the total number of bulls tested and number of cows mated per bull are varied, these numbers are not optimised. Also it should be mentioned that, in a practical setting, the costs that represent the cows used for progeny testing outside of the stud herd will prevail.

**5.4.1.3 Case C) As case B) but now using the LAMS approach,  
i.e. optimising the mean of grand progeny.**

In this scenario, unlike the previous cases, the prediction of the response on grand progeny was used as an objective function, instead of the calculated gain per year that is based on predicting offspring one generation ahead. In the previous section (no-LAMS) the distribution of EBVs was relevant as this was the selection criterion and the best predictor of the mean of progeny. In this section, for each cohort, the distributions of the true breeding value (TBV) conditional on the average EBV of the selected group need to be defined as this is relevant for the prediction of selection differential in the progeny. Hence, for a cohort  $i$ , i.e. the progeny of the selected parents in this generation, there is a mean EBV as the mean of the distribution of TBV, this is:

$$mean_i = \sum_{j=1}^{nr\ of\ malecohorts} \sum_{k=1}^{nr\ of\ femalecohorts} 0.5 * (EBVm_j + EBVf_k) \quad (5.8)$$

where  $EBVm$  and  $EBVf$  are the mean EBV of each cohort  $j$  of males and  $k$  of females selected as parents, respectively. The standard deviation of the true breeding value is determined by the PEV (prediction error variance, Figures 5.1 and 5.2) and the variance of the EBV distribution of each cohort. For males or females selected as parents this can be expressed as:

$$\sigma^2 EBVsel_i = (1 - k_i) * r_i^2 * \sigma_g^2 \quad (5.9)$$

where  $k$  is the variance reduction coefficient of the cohort  $i$  of males or females selected as parents,  $r_i^2$  is the squared accuracy of the cohort  $i$  of males or females selected as parents and  $\sigma_g^2$  is the genetic variance.

In consequence, the calculation of the SD of the true breeding value of each cohort  $i$  (the progeny of the selected parents in this generation) is:

$$\sigma_{\text{cohort } i} = \sqrt{\sum_{j=1}^{\text{nr of malecohorts}} \sigma^2 EBVs_{selm_j} + \sum_{k=1}^{\text{nr of femalecohorts}} \sigma^2 EBVs_{self_k} + \sigma_g^2 ms} \quad (5.10)$$

where  $\sigma^2 EBVs_{selm_j}$  and  $\sigma^2 EBVs_{self_k}$  are the variance of the EBV distribution within each cohort  $j$  of males and  $k$  of females selected as parents respectively and  $\sigma^2 ms$  is the remaining variance of the parent cohorts which is equal to:

$$\sigma^2 ms = \sum_{i=1}^{\text{nr of cohorts}} (\sigma^2 B_i + \sigma^2 W_i) \quad (5.11)$$

where  $\sigma^2 B_i$  is due to the error variance of the predicted parent cohorts  $i$  EBV mean and  $\sigma^2 W_i$  is due to within parent cohorts  $i$  additive genetic variation, which in turn are equal to:

$$\sigma^2 B_i = \sum_{i=1}^{\text{nr of cohorts}} 0.25 * (PEVm_i + PEVf_i) \quad (5.12)$$

$$\sigma^2 W_i = \sum_{i=1}^{\text{nr of cohorts}} 0.25 * (2 - Fm_i - Ff_i) * \sigma_g^2 \quad (5.13)$$

where  $PEVm$  and  $PEVf$  are the prediction error variance of EBVs of sires and dams, respectively, and  $Fm$  and  $Ff$  are the average inbreeding coefficients of selected parents within each cohort and  $\sigma_g^2$  is the genetic variance. But, as inbreeding is not considered in the analysis of this chapter, then the SD of the true breeding value of the progeny born from matings between sires from age class  $j$  and dams selected from age class  $k$  is calculated as:

$$\sigma_{\text{cohort } i} = \sqrt{\sum_{j=1}^{\text{nr of malecohorts}} \sigma^2 \text{EBVsel}m_j + \sum_{k=1}^{\text{nr of femalecohorts}} \sigma^2 \text{EBVsel}f_k + 0.25 * (\text{PEV}m_j + \text{PEV}f_k) + 0.5 * \sigma_g^2} \quad (5.14)$$

As already mentioned, the prediction of the response in grand progeny was used as the objective function for LAMS. LAMS aims to maximize this function, i.e. aims to maximize the predicted mean of grand progeny. So for LAMS, the differential evolution algorithm used in the macro was modified to optimise the value for the predicted selection differential from selection in the progeny of the next generation, which is the mean of selection among those progeny, or, in other words, the predicted mean of grand progeny. The maximum value for this last mean is achieved by trialling values for  $p$  of males and females along age classes (i.e the number of males and females selected from the corresponding age class as a function of the total number of males and females selected). Thus, there are 16 cohorts from combining 4 sire age classes with 4 dam age classes selected in the progeny of the next generation. Each cohort has a mean EBV based on the mean EBV of the parents. Also, each cohort has a distribution of EBV. The mean of the grand progeny is found by truncation selection across all these 16 cohorts, and the mean of the selected fractions or cohorts are equal to the cohort mean plus its selection differential. Then, the predicted mean of grand progeny is calculated as:

$$\text{cohortmeansel} = \sum_{i=1}^{\text{nr of fractions}} (\text{frac}_i * (cS_i + \text{frac}\mu_i)) \quad (5.15)$$

where  $\text{cohortmeansel}$  is the mean of selected fractions or cohorts of the progeny of the next generation,  $\text{frac}$  is the proportion of the selected individuals of progeny of fraction  $i$  out the total selected individuals of those progeny,  $cS$  is the selection differential of fraction  $i$  calculated as the product of the selection intensity of fraction  $i$  times the SD of EBVs of fraction  $i$  and  $\text{frac}\mu$  is the mean

of the selected individuals of fraction  $i$ . It should be noted that the SD of EBVs of fraction  $i$  is calculated as the product of the true genetic SD of that fraction or cohort (i.e.  $\sigma_{\text{cohort}_i}$ , equations [5.10] and [5.14]) times an assumed accuracy of that fraction or cohort. Note also that  $\text{frac}$  sums to 1.

The same initial set of data with low initial accuracy for the 2-year-old age class used in the previous case B (shown in Table 5.1 above) was used for no-LAMS and LAMS. For both no-LAMS and LAMS the difference between the predicted mean of age class 1 and age class 2 was used to measure the genetic progress for each year. The resulting proportions and number of proven and young bulls selected for the nucleus and the rate of genetic improvement are compared between the no-LAMS and LAMS approaches. In addition, as in cases A and B, the results of using 2 other alternative number of proven bulls as candidates are also analysed, i.e. 100 or 200.

## **5.5 Results**

### **5.5.1 Case A) Optimal proportions of selected sires in a herd with two age classes for males assuming differences in EBV accuracy between these age classes.**

The resulting relationships among genetic parameters of proven and young sires were derived from the Rendel and Robertson formula. Two examples are used in the following table to show these relationships.

In Table 5.2a, it can be observed that the proportion of selected proven or young sires increases as their respective (relative) accuracy increases, while the proportion of selected proven sires increases as the proportion of young sires decreases, as the sum of both proportions must add up to 1. A higher proportion of selected sires from the younger age class corresponds to a lower generation interval. In addition, lower selection differential of both proven and young sires

results as the accuracy of young bulls decreases. For the young bulls this is directly due to the lower accuracy and for the proven bulls it is due to the lower selection intensity.

Accuracy and mean of selected sires appear as the parameters with a major impact on the resulting proportions of selected young and proven bulls. The difference of accuracy between older and younger bulls age classes on the distribution of estimated breeding values as well as the mean of selected sires appears to have a major effect when selecting across age classes and consequently on the achieved genetic response.

Table 5.2a Examples with lower and higher accuracy for young bulls in the spreadsheet model developed for case ‘A’, where the parameter optimised is  $p$ : the proportion of bulls from each age class as proportion of total number of bulls used.

Parameters	Example 1		Example 2	
	Proven Bulls	Young Bulls	Proven Bulls	Young Bulls
Accuracy	0.80	0.40	0.80	0.30
Selected number	1.8	23.2	2.9	22.1
$p$	0.07	0.93	0.11	0.89
Mean of selected	1.53	1.34	1.35	1.08
Selection intensity	1.92	2.09	1.69	2.11
Selection differential	1.53	1.84	1.35	0.63
Response per year	0.254		0.225	

Results of the base scenario with 25 proven bulls (Table 5.2a, Example 1) are compared with 100 or 200 proven bulls as candidates, as shown in Table 5.2b. It can be seen that when maintaining the same accuracies of each age class, the proportion of selected proven sires increases as the number of progeny tested bulls increases. This allows higher selection intensities in both age classes which determines, in turn, the increase in the corresponding selection differentials and means, the genetic response and the average generation interval.

Table 5.2b Optimised parameters increasing the number of proven bulls candidates, maintaining the accuracies of each age class.

Parameters	100 proven bulls candidates		200 proven bulls candidates	
	Proven Bulls	Young Bulls	Proven Bulls	Young Bulls
Accuracy	0.80	0.40	0.80	0.40
Selected number	6.1	18.9	10.3	14.7
p	0.24	0.76	0.41	0.59
Mean of selected	1.58	1.40	1.64	1.46
Selection intensity	1.98	2.18	2.05	2.28
Selection differential	1.58	0.87	1.64	0.91
Response per year	0.264		0.275	

**5.5.2 Case B) Optimal selected proportions with respect to mean of progeny in a herd with several age classes for males and females with varying initial accuracy for the 2 year-old-age class of sires and the test capacity.**

In case B, the differential evolution algorithm tries to maximize an objective function, which is the mean of progeny in the next generation. The following Figures and Tables compare two situations: one with a low initial accuracy for the 2-year-old age class of sires (e.g. due to measuring few traits early on), and other with a high initial accuracy for the 2-year-old age class of sires (e.g. due to measuring more traits early on or using genomic information) (see above Table 5.1 to recall the sets of accuracies used for bulls and females in this case).

In Figure 5.8 the genetic gains achieved over time are shown. It should be noted that genetic gain becomes more stable, i.e. with less fluctuation between consecutive years, consistently with the proportions of selected sires along age classes, after around 10 mating seasons.

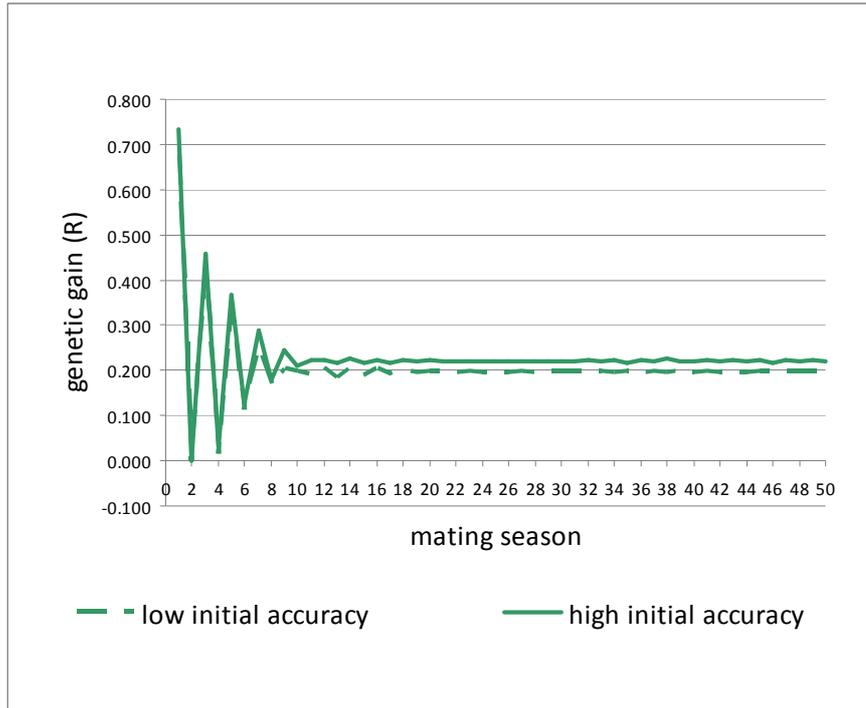


Figure 5.8 Genetic gains using low or high initial accuracy for the 2-year-old age class of sires.

Table 5.3a shows the equilibrium results in sires after running the breeding program for 50 mating rounds.

Table 5.3a Optimal selection results using low or high initial accuracy for the 2-year-old age class of sires.

Scenario		low initial accuracy			high initial accuracy		
Parameters		nr. bulls selected	accuracy	mean selected	nr. bulls selected	accuracy	mean selected
age class	2	22.2	0.37	9.85	23.3	0.43	11.04
	3	0.8	0.43	9.65	0.8	0.50	10.82
	4	0.5	0.50	9.45	0.0	0.56	10.59
	5	1.5	0.78	9.26	0.9	0.79	10.38
Avg L sires		2.25			2.14		
Response per year		0.197			0.221		
$\sigma_g$		0.91			0.90		

L refers to generation interval, Response per year to response per year after reaching equilibrium,  $\sigma_g$  to genetic variance.

Results in Table 5.3a show differences in the number of young (age classes 2 to 4) and proven bulls (age class 5) selected as sires along age classes. The number of selected young sires is increased when sires have high initial accuracy for the 2-year-old age class. Thus, in Table 5.3a appear that the total young sires from age class 2 to 4, i.e. 23.5 vs. 24.1, are the optimum numbers of selected sires for mating in these 2 situations respectively. On the other hand, the optimum numbers of proven bulls are 1.5 vs. 0.9 respectively. These represent 94.0 and 6.0% vs. 96.4 and 3.6% of young and proven selected sires respectively between the 2 analysed situations.

The numbers of selected females are also affected (see Table 5.3b), with fewer younger selected females (age classes 2 to 4) when the initial accuracy for the 2-year-old age class of sires is low (i.e. 894 vs. 909 respectively in the 2 situations). At the same time, the opposite occurs with the older females (age class 5), i.e. a bigger number are selected in the base scenario (106 vs. 91 respectively). These represent 89.4 and 10.6% vs. 90.9 and 9.1% of young and older selected females respectively between the 2 analysed situations. However, the differences are relatively smaller for selected females than for selected sires between both analysed situations.

Table 5.3b Optimal selection results for dams using low or high initial accuracy for the 2-year-old age class of sires.

Scenario		low initial accuracy			high initial accuracy		
Parameters		nr. dams selected	accuracy	mean selected	nr. dams selected	accuracy	mean selected
<b>age class</b>	<b>2</b>	411	0.31	11.35	427	0.29	11.04
	<b>3</b>	297	0.34	11.13	305	0.32	10.82
	<b>4</b>	186	0.37	10.91	177	0.35	10.59
	<b>5</b>	106	0.39	10.69	91	0.38	10.38
<b>Avg L dams</b>		2.99			2.93		
<b><math>\sigma_g</math></b>		0.91			0.90		
L refers to generation interval, $\sigma_g$ to genetic variance.							

It should be noted that, even though there are differences in mean between cohort age classes, the mean of selected parents within age class are similar, due to common threshold (see Figures 5.5b and 5.6 in section 5.4.1.2).

Higher genetic gains for the next generation are reached when the initial accuracy for 2-year-old age class of sires is higher compared with the case when this accuracy is low. These are 0.221 and 0.197 respectively (see Table 5.3a). The selection response increases mainly due to the higher mean of younger bulls selected as parents. Consistently with these results, a lower average generation interval in sires (2.14 and 2.25 respectively) and a slightly lower genetic variance in cohorts along all age classes (0.90 and 0.91 respectively) are observed (see also Table 5.3a).

In addition, in Figure 5.9 it can be observed that the genetic variance of the population is reduced by selection along the first 5 mating seasons due to the impact of the Bulmer effect, becoming more stable after that point.

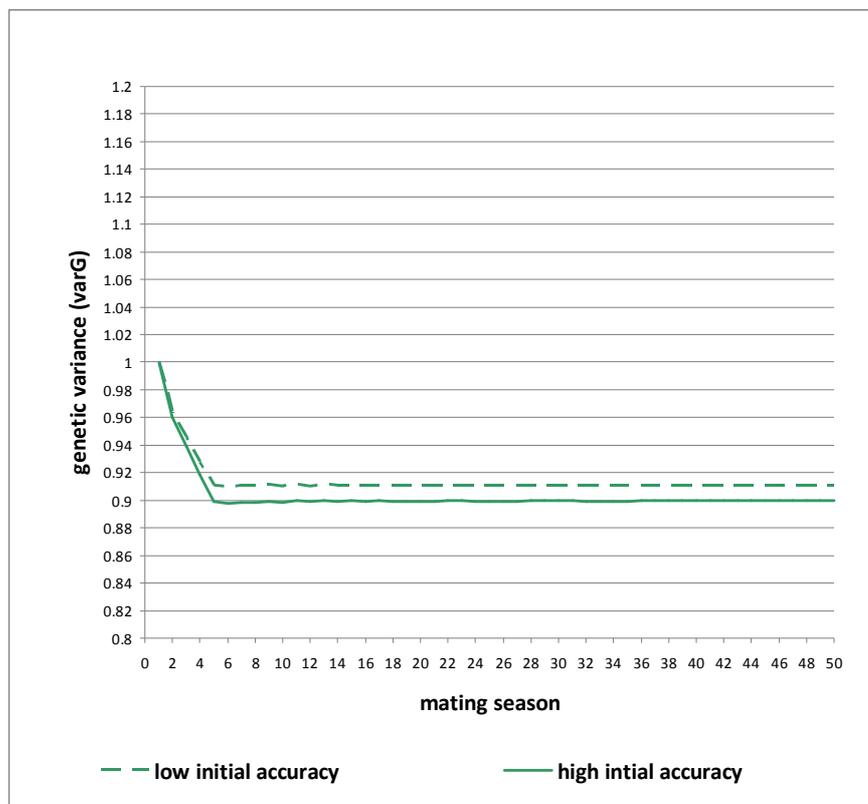


Figure 5.9 Population genetic variance using low or high initial accuracy for the 2-year-old age class of sires.

As mentioned previously in section 5.4.1.2, this reduction in variance in the new cohort of progeny born translates in a change in accuracy. This last fact can be noted comparing the resulting accuracies along age classes shown in Tables 5.3a and 5.3b with the initial accuracies used for bulls and females shown in Table 5.1.

In the following Tables the consequences of varying the test capacity are shown (i.e. the number of young bulls (age classes 2 to 4) that are progeny tested annually or progeny per young bull tested per mating season). Table 5.4a shows the consequences of the annual increase in the number of progeny tested 2-year-old bulls from 25 to 100 or 200 that allows later use of those numbers of 3 to 4-year-old bulls and proven bulls as candidates.

Table 5.4a Optimised parameters increasing the number of proven bulls candidates.

Scenario		100 proven bulls candidates			200 proven bulls candidates		
Parameters		nr. bulls selected	accuracy	mean selected	nr. bulls selected	accuracy	mean selected
<b>age class</b>	<b>2</b>	15.5	0.34	10.11	9.0	0.31	10.36
	<b>3</b>	2.4	0.41	9.91	3.2	0.39	10.16
	<b>4</b>	6.5	0.44	9.71	2.3	0.46	9.96
	<b>5</b>	5.6	0.77	9.51	10.5	0.77	9.75
<b>Avg L sires</b>		2.29			3.57		
<b>Response per year</b>		0.200			0.203		
<b><math>\sigma_g</math></b>		0.89			0.87		

L refers to generation interval, Response per year to response per year after reaching equilibrium,  $\sigma_g$  to genetic variance.

Comparing the results of the base scenario (see Table 5.3a on the left) to those in Table 5.4a, it can be observed that the proportions of selected 3 to 5-year-old sires increase while the proportion of selected 2-year-old sires decreases as the number of proven bull candidates increases. This determines, in turn, the increase in the means over age classes, the genetic response and the average generation interval. On the other hand, the accuracies across age classes of sires and the genetic variance decrease.

Then, the capacity for progeny testing is varied using a lower mating ratio, therefore testing more young sires in the nucleus. This implies not only a reduction of the number of dams mated per sire per young bull tested per mating season but also, at the same time, an increase in the total number of sires mated per year (age classes 2 to 5 of sires). Comparing the base scenario (40 dams per sire) to this new one (30 dams per sire), in Table 5.4b an increase is observed in the total number of sires mated per year from 25 to 33 respectively, (i.e. 32% more bulls mated). It should be noted that the total number of females of the herd has not been varied (i.e. it remains at 1000 females). It is also observed in this new scenario a slight proportional decrease in the number of young bulls selected (age classes 2 to 4 of sires) and, therefore, an increase of the number of selected proven bulls (age class 5 of sires), i.e. 31.7 and 2.3 respectively, which represent 93.0% and 7.0% of selected young and proven sires, respectively, compared to 94.0% and 6.0% of selected young and proven sires in the base scenario. Therefore, due to the proportionally lower number of selected young sires, a higher average generation interval resulted in sires (2.31 vs. 2.25 in the base scenario) even though a similar genetic variance existed in cohorts along all age classes (0.91 in both scenarios).

Table 5.4b Relevant optimal results varying the test capacity by modifying the used mating ratio.

Scenario		mating ratio = 40 dams per sire			mating ratio = 30 dams per sire		
Parameters		nr. bulls selected	accuracy	mean selected	nr. bulls selected	accuracy	mean selected
<b>age class</b>	<b>2</b>	22.2	0.37	9.85	28.6	0.36	9.44
	<b>3</b>	0.8	0.43	9.65	1.3	0.43	9.25
	<b>4</b>	0.5	0.50	9.45	0.8	0.49	9.06
	<b>5</b>	1.5	0.78	9.26	2.3	0.78	8.87
<b>Avg L sires</b>		2.25			2.31		
<b>Response per year</b>		0.197			0.188		
<b><math>\sigma_g</math></b>		0.91			0.91		
L refers to generation interval, $\sigma_g$ to genetic variance.							

Also, lower genetic gains for the next generation were achieved in this scenario (0.188 vs. 0.197 in the base scenario) due to lower genetic means and lower selection intensities along age classes. In addition, it is noted that accuracies along age classes do not change significantly between both scenarios.

As mentioned previously, the Bulmer effect was considered in cases B and C in the evolution of all parameters over the years of mating. The impact on the resulting parameters was examined if the Bulmer effect would not have been taken into account in the analysis. In Table 5.5 significant consequences of this omission are observed on the rate of gain while there is a small effect on optimal age structure of selected sires.

Table 5.5 Impact on results not considering the Bulmer effect using no-LAMS approach.

Approach		Bulmer effect			no-Bulmer effect		
Parameters		nr. bulls selected	accuracy	mean selected	nr. bulls selected	accuracy	Mean selected
age class	2	22.2	0.37	9.85	23.9	0.47	12.17
	3	0.8	0.43	9.65	0.7	0.52	11.92
	4	0.5	0.50	9.45	0.0	0.56	11.67
	5	1.5	0.78	9.26	0.4	0.75	11.43
Avg L sires		2.25			2.08		
Response per year		0.197			0.248		
$\sigma_g$		0.91			1.00		

L refers to generation interval, Response per year to response per year after reaching equilibrium,  $\sigma_g$  to genetic variance.

It is observed that the genetic means over age classes increase due to a lack of accounting for the Bulmer effect. It should be remembered that the genetic mean of the animals born each year is equal to the achieved genetic response in the parental generation. Hence, the predicted rate of genetic gain would be 25.9% too high if the Bulmer effect had been overlooked.

A more important consequence of not taking into account the Bulmer effect could be the observed change of the numbers of selected sires from different age classes, with relatively more young bulls selected. It appears that

the calculated genetic means affect selection relatively more than accuracies, causing more young sires to be selected, which in turn causes a lower average generation interval. In other words, the higher genetic response achieved per year results in a higher number of younger bulls selected as sires, after optimisation.

### 5.5.3 Case C) As case B) but now using the LAMS approach, i.e. optimising the mean of grand progeny.

The use of a selection criterion that uses the predicted response on grand progeny leads to a larger proportion of young candidates to be selected as sires which, as it was already mentioned, increases the genetic variance of progeny in the next generation. It should be remembered that, as in case B, the Bulmer effect was considered in the evolution of all parameters along the years of mating.

Table 5.6 Relevant optimal results using no-LAMS and LAMS approaches in the conditions of the base scenario

Scenario		no-LAMS			LAMS		
Parameters		nr. bulls selected	accuracy	mean selected	nr. bulls selected	accuracy	mean selected
age class	2	22.2	0.37	9.85	22.5	0.37	9.86
	3	0.8	0.43	9.65	0.9	0.44	9.67
	4	0.5	0.50	9.45	0.6	0.50	9.47
	5	1.5	0.78	9.26	1.0	0.78	9.27
Avg L sires		2.25			2.20		
Response per year		0.1968			0.1970		
$\sigma_g$		0.910			0.913		
L refers to generation interval, Response per year to response per year after reaching equilibrium, $\sigma_g$ to genetic variance.							

Comparing the results of these approaches for the base scenario where 25 young bulls are progeny tested (see Table 5.6) it is observed that the numbers of bulls selected as sires at different age classes at the 50th year are marginally different between no-LAMS and LAMS. In addition, in the case of LAMS, a slightly larger proportion selected from the young bull age classes is observed and a non-significant increase in rate of genetic gain.

Figure 5.10 (below) shows the difference of the population with LAMS and no-LAMS in the conditions of the base scenario.

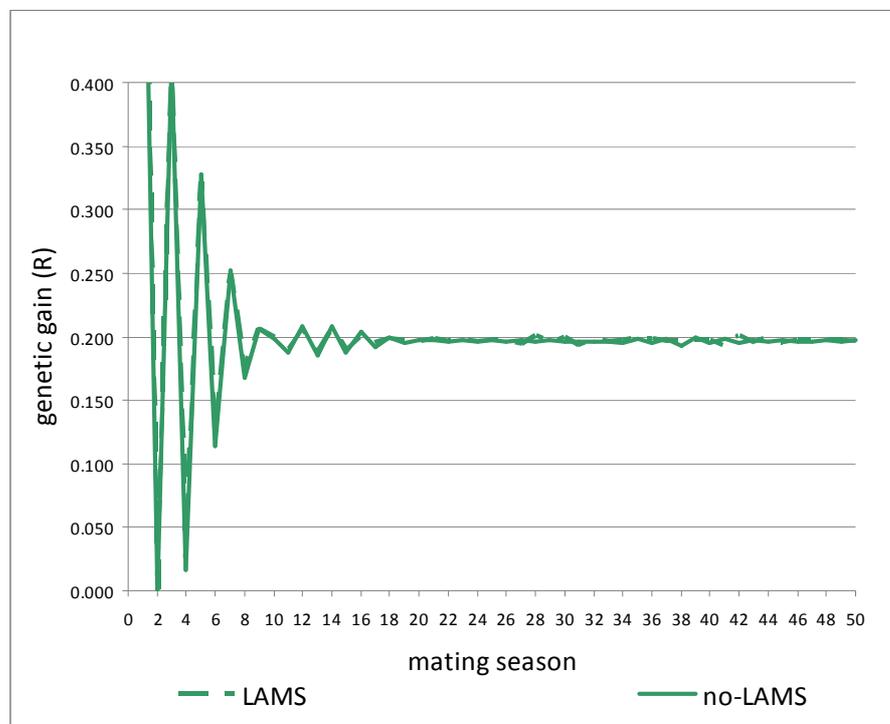


Figure 5.10 Genetic gains using no-LAMS and LAMS approaches in the conditions of the base scenario

Slightly higher genetic gains along years are observed in LAMS than in no-LAMS, i.e. very small differences appear between the obtained genetic gains in both approaches, almost indiscernible in Figure 5.10.

Table 5.7 Relevant optimal results using no-LAMS and LAMS approaches, with 100 proven bulls candidates.

Scenario		no-LAMS with 100 proven bulls candidates			LAMS with 100 proven bulls candidates		
Parameters		nr. bulls selected	accuracy	mean selected	nr. bulls selected	accuracy	mean selected
<b>age class</b>	<b>2</b>	15.5	0.34	10.11	17.8	0.36	10.33
	<b>3</b>	2.4	0.41	9.91	2.6	0.43	10.13
	<b>4</b>	1.6	0.44	9.71	1.6	0.49	9.92
	<b>5</b>	5.6	0.77	9.51	3.6	0.78	9.72
<b>Avg L sires</b>		2.89			2.60		
<b>Response per year</b>		0.200			0.206		
<b><math>\sigma_g</math></b>		0.89			0.90		
L refers to generation interval, Response per year to response per year after reaching equilibrium, $\sigma_g$ to genetic variance.							

Table 5.8 Relevant optimal results using no-LAMS and LAMS approaches, with 200 proven bulls candidates.

Scenario		no-LAMS with 200 proven bulls candidates			LAMS with 200 proven bulls candidates		
Parameters		nr. bulls selected	accuracy	mean selected	nr. bulls selected	accuracy	mean selected
<b>age class</b>	<b>2</b>	9.0	0.31	10.36	14.0	0.35	10.79
	<b>3</b>	3.2	0.39	10.16	3.5	0.42	10.57
	<b>4</b>	2.3	0.46	9.96	2.2	0.49	10.36
	<b>5</b>	10.5	0.77	9.75	5.3	0.77	10.15
<b>Avg L sires</b>		3.57			2.95		
<b>Response per year</b>		0.203			0.215		
<b><math>\sigma_g</math></b>		0.87			0.90		
L refers to generation interval, Response per year to response per year after reaching equilibrium, $\sigma_g$ to genetic variance.							

However, when comparing the results with scenarios where more young bulls are progeny tested, e.g. to have 100 or 200 proven sires available as candidates (see Tables 5.7 and 5.8), it can be observed that the differences between no-LAMS and LAMS increase moderately. Thus, the numbers of bulls

selected as sires at different age classes at the 50th year under no-LAMS and LAMS are moderately different, with, in the case of LAMS, a larger proportion selected from the young bulls age classes and a slightly higher rate of genetic gain and, therefore, slightly higher means after 50 rounds of selection.

Figure 5.11 (below) shows the difference of the population with LAMS and no-LAMS when 100 proven sires are available as candidates. Slightly higher genetic gains are observed over years in LAMS than in no-LAMS, i.e. moderate differences appear between the obtained genetic gains in both approaches, with LAMS giving around 3.0% more annual gain. It should be noted that the mentioned differences between LAMS and no-LAMS become even larger as larger is the number of proven sires to have available as candidates, with LAMS giving around 5.9% more annual gain with 200 proven sires and around 8.6% more annual gain with 450 proven sires.

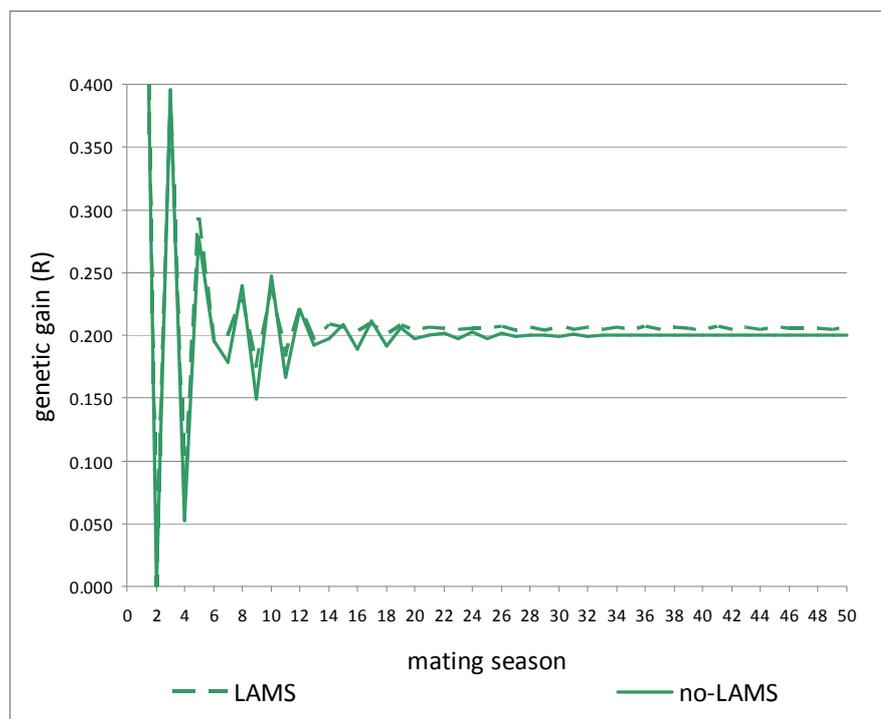


Figure 5.11 Genetic gains using no-LAMS and LAMS approaches, with 100 proven bulls candidates.

Consistent with the other results shown in Table 5.8, it is observed that the average generation interval (calculated as the average between the generation interval of selected males and females in each year) of LAMS is lower than the average generation interval of no-LAMS (see previous Table 5.8 and Figure 5.12). The reduction of generation interval is associated with the acceleration of the rate of genetic gain.

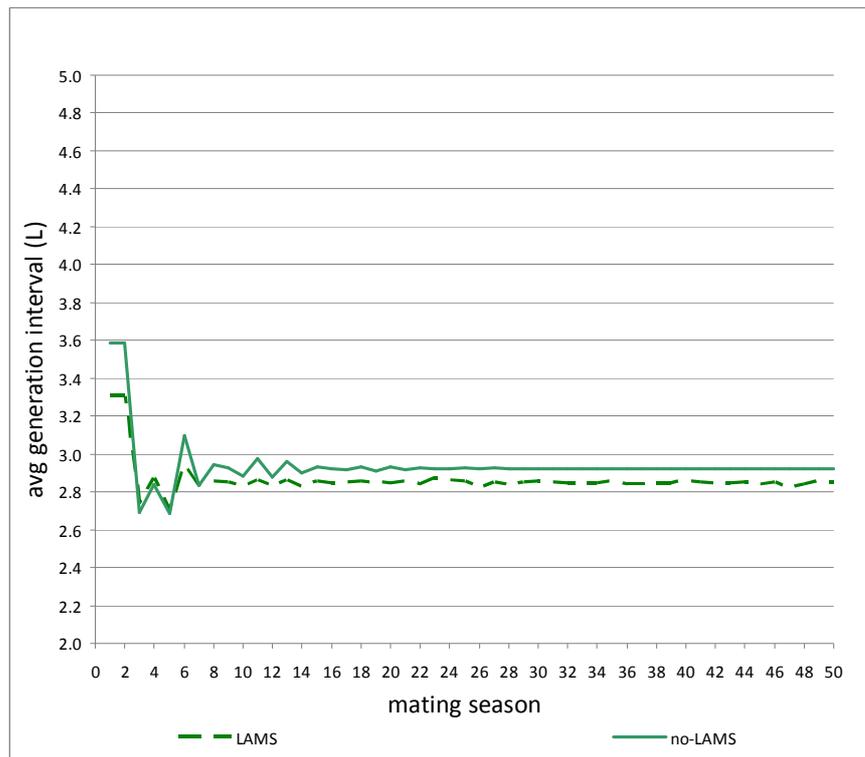


Figure 5.12 Average generation intervals using no-LAMS and LAMS approaches with 100 proven bulls candidates.

## 5.6 Discussion

The results show in LAMS a larger proportion of selected sires from the young bull age classes and a slightly higher rate of genetic gain and, therefore, slightly higher means after 50 rounds of selection. This higher genetic gain increases with more sire candidates in the older age classes. Around 3.0% more genetic gain is obtained in LAMS compared with no-LAMS when 100 proven

bulls are available as candidates to be selected as sires. However, it is necessary to take into account that the cost to obtain larger numbers of proven sires would increase considerably.

Through a Progeny Testing program the genetic merit of an animal in the population is tested for selective breeding by looking at its progeny produced by different matings. Progeny testing is most useful and commercially important to determine the true breeding value of an animal, i.e. when a high level of accuracy is needed for selecting a sire to be used extensively in artificial insemination. The extensive use of artificial insemination in beef cattle has helped to increase the selection intensity among bulls. A larger number of progeny from an individual eliminates some of the non-genetic sources of error and increases selection accuracy. Also, evaluating an animal on the basis of progeny performance tends to eliminate error from non-additive gene action. A disadvantage of progeny testing is that only a very limited number of individuals can be progeny tested compared to the total number produced. This selection tool is usually used for traits that are sex-limited, expressed after death (like e.g. tenderness of meat) and also with low heritability. However, progeny testing many more young bulls is becoming uneconomical as young bulls that are progeny tested are having a low chance to ever be selected as proven sires.

Progeny testing could be conducted in a large herd or involving associated herds or in the field in farmers place. A large nucleus, in this case a dispersed nucleus which comes from several stud-farms with a limited herd size, has the advantages in a YSP that provides an opportunity to progeny test more young bulls in the nucleus, which contributes to genetic progress in the nucleus with the use of these best young sires and which would be an expensive exercise for individual studs. Also, this allows breeders to share risk and opportunity by testing progeny of young bulls in each others herds, which also links up herds very well. This is in turn an important aspect of across herd sire evaluation and could improve the rate of genetic gain with less increase in inbreeding. A smaller nucleus (of around 500 dams) could also be used, maybe still dispersed

over many studs, not all cows being in the stud and mated with elite sires. A field based progeny test system is required when the selected bulls are to be distributed in a large area, to many studs in different environments. Conducting this last kind of progeny testing, especially in small holder production systems, requires resources both financial and infrastructural (i.e. a large AI network, robust and dynamic data collection and analysis system).

Thus, joining a YSP appears as a suitable solution for stud-farms with a limited herd size that exists in Argentina, decreasing the cost of genetic progress compared with the use of imported semen, and commercial clients benefit from more accurate breeding values and better genetics.

As already mentioned, one of the main objectives of a progeny test is to accurately test the young sires to use in a breeding program. The achievement of this goal allows, in turn, higher genetic gains along selection rounds, which is one of the main objectives of a breeding program. After optimizing using young sires as nucleus sires, a YSP would emerge (i.e. using many young sires as nucleus sires) if that is the best solution. All this is shown through the results obtained in this study where, in a fixed sized total herd, this aim is achieved by increasing the proportion of selected young bulls as sires rather than increasing the number of young sires as a result of increasing the total amount of selected sires.

Banks et al. (1998) have examined the optimal proportions of young and old sires. Meszaros et al. (1999) used for this purpose a genetic algorithm. A comparison between the results of this study and the mentioned works appears not feasible due to the different method of optimisation, number of age classes and population size used.

There seems to be little or no previous work similar to the analysis carried out in this Chapter, probably because of a relative lack of development of mate selection systems that operate beyond control of genetic merit and coancestry.

In this study the equations derived from the Rendel and Robertson formula were used to predict genetic gain, for a given set of proportions selected

from each age class. Other genetic parameters such as selection accuracy and generation interval resulted from that. The use of a DE algorithm was the key tool to determine the optimal proportions of proven and young bulls to be selected as parents along age classes and balance genetic parameters.

The application of truncation selection based on BLUP EBV to optimise the selection of sires in YSP, i.e. to reserve the optimal proportion of matings for young and proven sires, is a strategic optimisation, involving simply selecting the sires with the best breeding value. For this optimisation it was necessary to set a whole population model of the breeding program in which the different parameters were accommodated. Thus, breeding alternatives were compared by varying those parameters from which different strategies were derived. Meanwhile, the LAMS selection method looks for selecting relatively more young sires to capitalise on the higher residual variation in genetic merit that they have. This strategy would result in selecting more young sires as nucleus sires with less accurate estimated breeding value but with more prediction error variance. This is achieved by maximising the predicted genetic merit in the grand progeny considering merit two generations ahead (see above equation [5.15]) through the use of a DE algorithm. This sets up appropriate reward for the increased future selection opportunity given by choosing less accurately evaluated animals, which is especially valuable under overlapping generations (Goddard and Howarth, 1994). Thus, LAMS can be implemented as a tactical optimisation because the higher PEV for some individuals can be predicted and exploited in a competitive manner to give higher future selection differentials and more genetic gain in the long term.

It also should be noted that the LAMS approaches used in Chapter 3, on one hand, and Chapters 5 and 6, on the other, differ (see this contrast in section 1.2).

Inbreeding was not accounted for in the analysis. This could be considered as a shortcoming because inbreeding reduces the genetic variance. Not accounting for inbreeding causes an overestimation of the achieved genetic

gains. It should be noted that inbreeding depends on nucleus size, and that it can be controlled by optimal contribution selection, as will be done in Chapter 6.

Considering the consequences of not taking into account the Bulmer effect in this analysis, it could be added that, in reality, earlier breeding values are more determined by family information, and it is the between family variation that is reduced by the Bulmer effect. As this was a deterministic simulation, correlation of EBVs among relatives was not accounted for. It would have needed a stochastic simulation to account for that and for inbreeding. That correlation and inbreeding will be considered in the analysis in Chapter 6.

## ***5.7 Conclusions***

Optimal proportions and number of proven and young bulls and females selected for the nucleus are achieved by optimising a suitable objective function. This optimisation used Differential Evolution algorithm by trialling values for the proportions of males and females selected out of the total number of males and females of their corresponding age classes, and evaluating these values by the prevailing objective function. This objective function is, in the case of no-LAMS (cases A and B), the selection differential of the selected males and females as parents of the next generation or, in the case of LAMS (case C), the predicted genetic merit in the grand progeny, i.e. two generations ahead.

The LAMS approach results in a slightly higher genetic response compared to no-LAMS. This marginally higher rate of genetic gain increases more as more young bulls are progeny tested, increasing in this way the number of sire candidates across all age classes. Thus, around 3.0% more genetic gain is achieved in LAMS than in no-LAMS when 100 proven bulls are available as candidates to be selected as sires. In addition, the distribution of elite sires over age classes is affected, with more sires selected from the younger age classes and a slightly higher rate of genetic gain.

The higher the difference in accuracy between age classes, the more the no-LAMS approach will tend to increase the number and proportion of selected animals in the age class with higher accuracy, while the LAMS approach will tend to increase the number and proportion of selected animals in the age class with lower accuracy. In other words, a higher increase in the number and proportion of selected animals in the age class with higher accuracy will occur in no-LAMS approach, while in LAMS approach it will be higher in the age class with lower accuracy.

## **CHAPTER 6**

# **STOCHASTIC EVALUATION OF LOOK AHEAD MATE SELECTION FOR A YOUNG SIRE PROGRAM**

## **6.1 Summary**

In this Chapter the consequences of applying a Look-Ahead Mate Selection (LAMS) scheme approach are analysed in a Young Sire Program (YSP) scenario through a stochastic simulation. In LAMS, the aim is the maximization of the rate of genetic progress along the years of mating by the maximization of the predicted genetic selection differential in the next generation instead of maximizing the progeny additive genetic merit in the next generation as in the traditional scheme (no-LAMS). This is achieved by using proper objective functions that are managed by a differential evolution algorithm. The main question to answer will be whether tactical Mate Selection can in fact set up such a scheme, and whether it improves longer-term response by making opportunistic use of the prevailing animal resources.

Due to selection, the resulting average response in TBV was 0.289 for LAMS and 0.270 for no-LAMS, i.e. 7.04% higher for LAMS than for no-LAMS. Also it appears that the achieved TBV mean starts to be slightly higher in LAMS than in no-LAMS at the 14th year of mating ( $\approx 0.303\%$  at the 20<sup>th</sup> year of mating) but not significantly (p-value = 0.811).

The marginal superiority of LAMS appears only when the number of candidates to select from is sufficiently high in the herd. Otherwise, LAMS is not competitive. However, these comments should not be taken to imply that a sufficiently large herd size would necessarily make LAMS usefully better than no-LAMS.

Comparing these results to those in Chapter 5, in which the setup of a YSP was analysed through a deterministic analysis of modelling and simulation, it is observed that the resulting patterns of the uses of selected sires in this Chapter become closer to the observed pattern in Chapter 5 as the chosen target proportion of selection is lowered.

Thus, even though there were small differences between LAMS and no-LAMS, tending towards setup of a YSP under LAMS, the impact is very slight compared to imposing a designed YSP directly.

In conclusion, tactical Mate Selection can in fact set up a scheme that tends towards the features of a YSP.

## 6.2 Introduction

In general, breeders look for maximizing the progeny additive genetic merit in the next generation. For this, they select males and females as parents based on their EBVs. Frequently, as was analysed in previous chapters, breeders can consider other additional aspects during the selection and mating process, such as reduction of inbreeding and coancestry, corrective mating or the advantages of assortative mating.

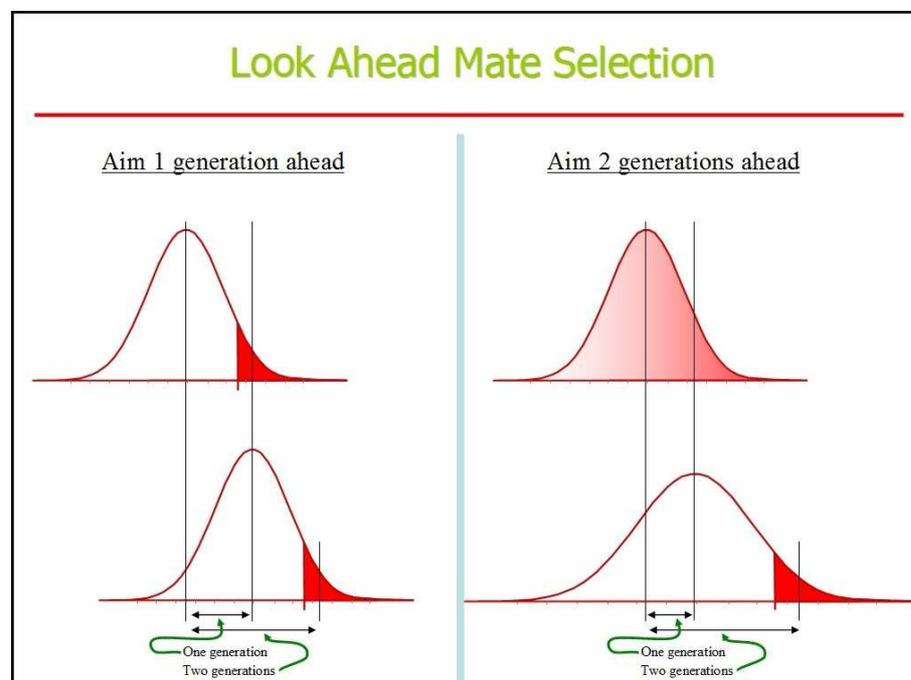


Figure 6.1 Look Ahead Mate Selection scheme.  
(Courtesy Prof. Brian Kinghorn)

As it was mentioned in section 5.2 of Chapter 5, Kinghorn and Shepherd (1994) have proposed an alternative approach, i.e. Look-Ahead Mate Selection (LAMS), in which the aim is the maximization of the predicted genetic selection differential in the next generation (see above Figure 6.1). The selection differential in the progeny generation can be predicted by the difference between the average of the entire population of prospective future progeny and the predicted average of the parents selected from them. Note that the evaluation of individuals is not delayed by a generation, but a prediction of second generation impact is immediately available such there is no deleterious impact via longer generation intervals.

Objective functions which aim to maximize the predicted genetic merit of the grand progeny of selected parents exploit positive assortative mating. This is improved by mate allocations which yield better information from the pedigree structure and benefits from connectedness of relatives in different fixed effects groups, e.g. herd or management group, allowing the better comparison of animals on EBV. In this scenario, two types of mating can be described:

- Investment matings
- Realisation matings

In our case, investment matings would be, in a Young Sire Program scenario, the use of young sires with a lower average EBV and lower accuracy and realisation matings the use of proven sires with a higher average EBV and higher accuracy. The former is suboptimal in the short term in order to invest for the long term, i.e. to realise the benefits in the latter by revealing residual genetic variation among young sires.

Thus, in the process of Mate Selection some weight is implicitly given to investment matings, which compete with realisation matings when maximizing predicted grand progeny merit. All this may reduce selection differential in the parent generation, as a consequence of selection of animals which would not be selected on EBV alone but can contribute sufficiently to a mating design which takes advantage of that better information. In this way, variance in predicted

genetic merit in the progeny generation is increased and this will generally be reflected by increased variance in true genetic merit. Even though a reduction in one-generation response is likely, this results in an increase in two-generation response, i.e. an increment of genetic merit in the second generation.

An extension of this approach was proposed by Segelke et al. (2014) in genomic selection in dairy cattle using specific haplotype information. Thus, parents were identified for mating based on the predicted genetic variation within groups of offspring for different traits, i.e. through the standard deviation of gamete breeding values (SDGBV) between animals at the haplotype level. The SDGBV depends on the variance of the true effects of single nucleotide polymorphisms (SNPs) and the degree of heterozygosity. So, sires that had relatively low mean gamete breeding values but high SDGBV had a higher probability of producing extremely positive offspring than sires that had a high mean gamete breeding value and low SDGBV.

In this Chapter the consequences of applying a Look-Ahead Mate Selection scheme approach in a Young Sire Program (YSP) scenario are tested and discussed through a stochastic simulation. The key question here will be whether tactical Mate Selection can in fact set up such a scheme, and whether it improves longer-term response by making opportunistic use of the prevailing animal resources.

### ***6.3 Objective of the analysis***

Two treatments or approaches were tested: no-LAMS and LAMS. As mentioned in Chapter 5, the former aims to maximize genetic response in the next generation, while the latter aims to maximize genetic response on grand progeny.

The method for LAMS was to select on the predicted genetic selection differential in the next generation as a deviation from current population mean,

giving the next generation plus one generation ahead response. Thus, the prediction of the genetic merit in grand progeny was used as the objective function for LAMS. For no-LAMS, the method was to select on merit of prospective progeny. Both methods employed optimal contributions, with a tendency for higher indexing males to be used in more matings.

For LAMS, the differential evolution algorithm maximises the value for the predicted genetic selection differential from selection in the progeny of the next generation, which is the predicted mean of grand progeny.

## ***6.4 Materials and methods***

### **6.4.1 Data simulation**

A base dataset of a seed stock herd of beef cattle with overlapping generations with 100 sires and 800 dams was generated using PopSim (version 1.5) which includes the Matesel system. The mating ratio averaged 8, but was variable because of optimal contributions, and each dam had one progeny. Foundation animals were assumed to be unrelated, unselected and randomly sampled from an infinite population. The phenotypic values of all the individuals in the population were simulated. Heritability was set low, at 0.05, to give enough difference in accuracy between own information and progeny testing information, and thus for older sires to benefit in their accuracy. This was done in order to help demonstrate any value of LAMS, rather than to make an evaluation of likely value under real conditions. The EBV and corresponding accuracy for all these individuals in the population were estimated using the BLUPF90 program of Ignacy Misztal (1999), in which a subroutine was added, using BLUPF90 features and facilities, to calculate the accuracies of EBVs from the diagonal elements of the inverse of the coefficient matrix of the Mixed

Model Equations, which is based on the method proposed by Henderson, 1984 (Monti A., 2004).

First progeny were born when parents were age 3 years; culling for males and females was handled by the Mate Selection program, i.e. the best sires and females were kept in the population for repeat mating compared to those with lower genetic merit; random adult annual survival of 95% was set. Under no-LAMS selection was on EBV, and under LAMS selection was on a function including EBV, selection accuracy, inbreeding coefficients and selection proportion, as described below. In this way, a population with overlapping generations according the selection criteria was obtained and its final genetic parameters could be estimated.

As mentioned, two treatments were tested: LAMS and no-LAMS.

The population simulation program was run to get 10 replicates (each one using a different seed number to avoid a sampling covariance) for each of the two treatments, i.e. no-LAMS and LAMS, and each treatment for 20 annual selection rounds.

It should be mentioned that the Bulmer effect is considered by the PopSim program.

In Matesel, Balance Strategy was set to 3 and target degrees were set to zero in order to maximize the predicted merit in the next generation, while a negative weighting of -1 on progeny inbreeding was imposed to achieve reduced inbreeding in the next generation (see section 3.3.3 of Chapter 3).

After the last mating season the simulated population had a total of 3265 live animals available for reproduction (1313 males and 1952 females, being fewer males because of fewer were selected into older age groups), each one with its own predicted genetic merit expressed in EBVs. The target proportion of selected parents (i.e. ' $p$ ') was set at 40%. Based on their genetic merit and corresponding accuracy, males and females of those available for reproduction were considered as candidates for selection, with a requirement to make 800

matings. Females had only 1 mating allowed as embryo transfer was not considered.

#### **6.4.2 Analysis method**

Following this structure, the Matesel option in the Popsim v1.5 system was then applied to determine the set of matings (including optimum contribution of candidates) to generate the next set of progeny. The source code of the Popsim system was modified by the author in order to achieve the custom analyses required in this chapter.

Chapter 5 applied a deterministic simulation to predict the genetic selection differentials, in which animals were sampled from several cohorts of sire and dam age classes combinations, and compared the results of maximising next generation response vs. when the selection response was achieved two generations ahead.

Now, in this Chapter, a LAMS scheme approach was developed but, unlike Chapter 5, through a stochastic simulation. Thus, in this analysis, the correlation of EBVs among relatives and inbreeding are accounted for. As this is a stochastic simulation, the criterion for YSP operates through predicted selection among progeny, rather than selection among the same candidates at a later stage in their life, as in Chapter 5.

As it was already mentioned, simulations were done over multiple selection rounds, being the approach fully tactical as the solution was optimised at every selection round.

Genetic selection differential in the current generation (i.e. at generation 0), which reflects genetic merit in generation 1, is calculated as the average of the means of selected sires and dams deviated from the mean of the candidates (or the average of the two sex means of candidates in case there is a systematic difference in EBV between male and female candidates).

Then, in order to calculate the selection differential one generation ahead (i.e. at generation 1, reflecting genetic merit in generation 2) it is required to set the progeny distribution in which the overall truncation point is determined to give the established target proportion 'p' selected across the distributions of matings.

The mentioned progeny distribution is a smooth 'infinite' population of distributions, which is the sum of the normal distributions of progeny from each mating. It should be noted that the distribution of a mating represents the possible outcomes of a mating between two individuals, giving a density distribution of merit for resulting progeny, no matter how many progeny arise per mating.

For each mating the mean and variance of the progeny distribution are calculated. This mean (i.e. the predicted genetic merit of the progeny group for each mating  $j$ ) is equal to the mean of the EBV of the two selected parents, while the variance is equal to the sum of two variances: one due to the error variance of the predicted family EBV mean [ $\sigma^2 B(j)$ ] and the other due to within family additive genetic variation [ $\sigma^2 W(j)$ ]. Hence, for a progeny group  $j$ , i.e. the progeny of the selected parents of the mating  $j$ , there is a mean EBV as the mean of the distribution of TBV, this is:

$$mean_j = 0.5 * (EBVm_j + EBVf_j) \quad (6.1)$$

where  $EBVm$  and  $EBVf$  are the EBV of the sire and dam selected as parents of the mating  $j$ , respectively. On the other hand, the calculation of the variance of the true breeding value of the progeny group of the selected parents of the mating  $j$  is:

$$\sigma^2_{mating_j} = \sigma^2 B_j + \sigma^2 W_j \quad (6.2)$$

where  $\sigma^2 B_j$  is the between family variance due to the error variance of the predicted parents' EBV mean of mating  $j$  and  $\sigma^2 W_j$  is the within family variance due to within parents' additive genetic variation of mating  $j$  which in turn are equal to:

$$\sigma^2 B_j = 0.25 * (PEVm_j + PEVf_j) \quad (6.3)$$

$$\sigma^2 W_j = 0.25 * (2 - Fm_j - Ff_j) * \sigma_g^2 \quad (6.4)$$

where  $PEVm$  and  $PEVf$  are the prediction error variance of EBVs of the selected sire and dam of mating  $j$ , respectively, and  $Fm$  and  $Ff$  are the inbreeding coefficients of selected parents of mating  $j$ , and  $\sigma_g^2$  is the population additive genetic variance. Then, the SD of the true breeding value of the progeny born from the mating between the selected sire and dam of mating  $j$  as a deviation from the predicted mating mean is calculated as:

$$\sigma_{mating_j} = \sqrt{0.25 * (PEVm_j + PEVf_j) + 0.25 * (2 - Fm_j - Ff_j) * \sigma_g^2} \quad (6.5)$$

As mentioned, an overall truncation point ( $X$ ) is necessary to calculate the selection differential at generation 1. The same value was used for males and for females, for simplicity. Different approaches through iterative algorithms can be used to determine  $X$ . One of these is the numerical summation approach, in which the proportion of mating distributions coincident to the target proportion of selected parents (' $p$ ') is calculated after obtaining the total area by the sum of heights (from the normal distribution height equation) across all distributions of matings, i.e. using deterministic sampling in an essentially perfect normal shape. Another approach (probably faster and more accurate) is using order statistics through the predicted progeny distribution variance and approximate relationships between the selection intensity ' $i$ ' and selected proportion ' $p$ '.

Also, a stochastic simulation of individual progeny could be used, which is less convenient due to it introduces unwanted random variation in the distribution shape. The second mentioned approach was chosen to apply, i.e. order statistics.

The distribution of means of matings is on an EBV scale (i.e. a predicted progeny genetic merit scale). This scale is on both sides of zero (which is the mean across the distributions), so some of them are positive, others negative. The full range used for finding  $X$  was taken as 4 SD of progeny merit above the mean of the biggest predicted progeny genetic merit value minus 4 SD of progeny merit below the mean of the smallest one (see Figure 6.2).

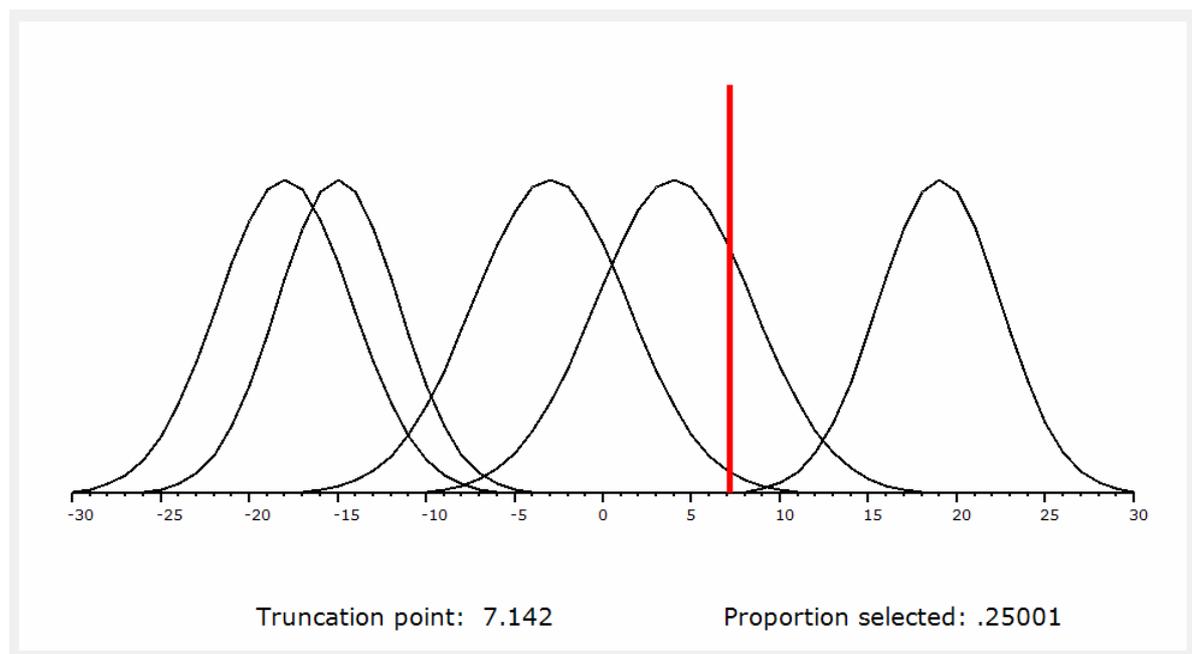


Figure 6.2 Diagram of scheme used to determine the overall truncation point  $X$  (red line) to give a target proportion selected (0.25 in this case) across the distributions of matings.

Setting all these parameters, the mentioned overall truncation point  $X$  is determined. The correct value is that by which the sum of the selected proportions or fractions for all the mating distributions, corresponding to truncation at  $X$ , is coincident to the target proportion of selected parents ' $p$ '. So,

the DE algorithm finds the optimal mating distributions (i.e. the selected parents at generation 1) to be “included” in the target proportion ‘ $p$ ’.

When the correct truncation point  $X$  is found, then the selection intensity for each mating distribution can be calculated.

Finally, at the end of this process, there will be a series of overlapping normal mating distributions with different means and variances, plus the ‘optimal’ truncation point and the corresponding selection intensities that are used to give the desired result, i.e. the required overall selection differential at generation 1 or the predicted mean of grand progeny at generation 2.

Thus, the predicted mean of grand progeny is the mean of the selected proportions or fractions of these mating distributions, which are equal, in turn, to the mean of the mating distribution plus its selection differential. Then, the predicted mean of grand progeny is calculated as:

$$cohortmeansel = \sum_{j=1}^{nr\ of\ matings} (prmat_j * (cS_j + fracmu_j)) \quad (6.6)$$

where *cohortmeansel* is the mean of selected mating distributions or fractions of the progeny of the next generation, *prmat* is the selected proportion of the mating distribution or fraction  $j$ ,  $cS$  is the selection differential of fraction  $j$  calculated as the product of the selection intensity of fraction  $j$  times the SD of EBVs of fraction  $j$ , and *fracmu* is the mean of fraction  $j$ . It should be noted that the SD of EBVs of fraction  $j$  is calculated as the product of the true genetic SD of that fraction (i.e.  $\sigma\ mating_j$ , equation [6.5]) times an assumed accuracy of that fraction. Also, it should be noted that the selected proportion of the mating distribution  $j$  (*prmat<sub>j</sub>*) is needed in this equation [6.6] because lower merit and thus more highly selected families contribute less to the predicted mean of grand progeny as high selection intensities mean fewer selected animals. On the other hand, the selected proportions of all mating distributions (i.e. all *prmat<sub>j</sub>*) must add to unity.

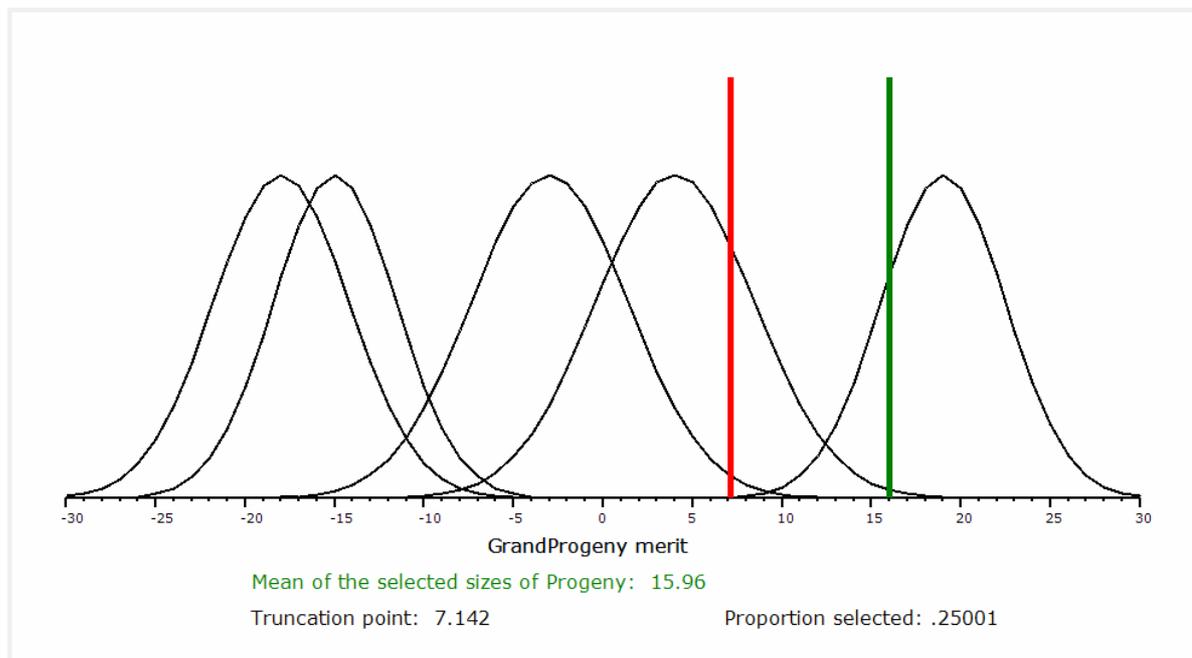


Figure 6.3 Diagram of scheme used to determine the predicted mean of grand progeny (green line) according a target proportion selected across the distributions of matings and the overall truncation point X (red line).

Applying all what was just mentioned to the diagram shown in Figure 6.2, Figure 6.3 (see above) was obtained.

The equations above are equal or similar to those used in Chapter 5 for LAMS. However, this Chapter mating distributions or families have been managed instead of cohorts across age classes.

In order to test whether extra response in “generation 2” is seen, i.e. to evidence the impact of LAMS in the selection response compared to no-LAMS, or, in other words, to test if the difference between the responses obtained following LAMS and no-LAMS schemes is significant, a one-tailed Z-test for independent samples was used. For this, it is necessary to handle many replicates to get a significant difference between these two strategies or schemes, taking into account the SD of that difference. As previously mentioned, it was decided to do runs considering 20 annual selection rounds (mating seasons) for each replicate. The results of 10 replicates for each of the treatments (i.e. LAMS and no-LAMS) were used for analysis.

Mean true breeding value (TBV) was used to measure selection response for each treatment (LAMS and no-LAMS) for each year of mating across replicates. Mean TBV is simpler to use, as the mean phenotype adds error ( $\sigma_e^2$ ) and the mean EBV depends on BLUP.

For a Z-test, is needed to calculate:

$$SE_{(tr)} = \sigma_{(tr)} / \sqrt{nreps_{(tr)}}$$

and

$$Z = Diff / SE \text{ for difference}$$

being

$$SE \text{ for difference} = \sqrt{SE_{(no-LAMS)}^2 + SE_{(LAMS)}^2}$$

where  $tr$  is a treatment,  $SE_{(tr)}$  is the standard error of the mean of the a treatment in selection response,  $\sigma$  is the standard deviation across replicates for a treatment,  $nreps$  is the number of replicates used in a treatment,  $Diff$  is the mean of the difference between treatments in selection response and  $SE \text{ for difference}$  is the standard error of the mean of the difference between treatments in selection response.  $Z$  is then used to test the significance of difference between LAMS and no-LAMS.

## 6.5 Results

The genetic gains achieved due to selection cause an increase in TBV mean along the years of mating for LAMS and no-LAMS as shown in Figure 6.4.

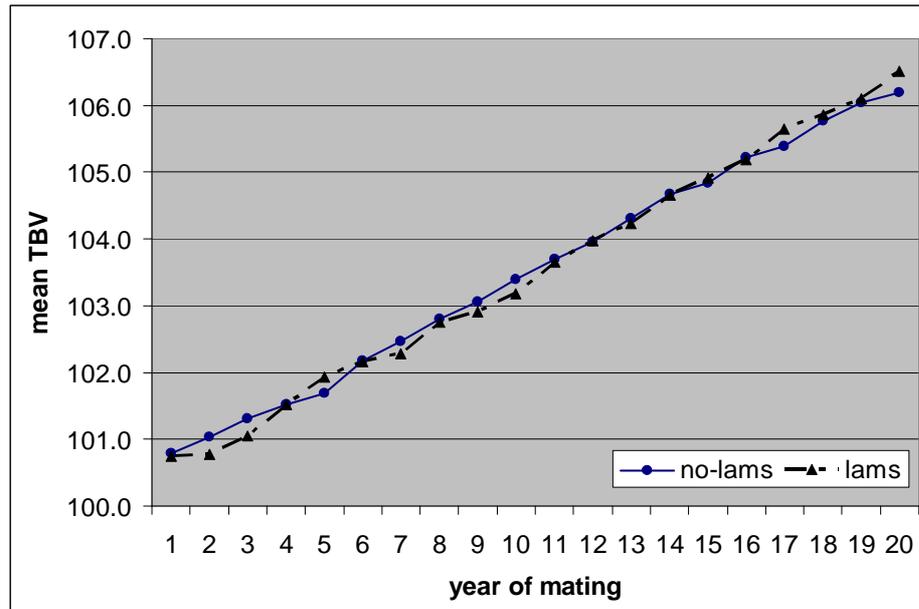


Figure 6.4 Trend of the TBV mean along the years of mating or LAMS and no-LAMS

The average response in TBV was 0.289 for LAMS and 0.270 for no-LAMS, i.e. 7.04% higher for LAMS than for no-LAMS. Also it appears that the achieved TBV mean starts to be slightly higher in LAMS than in no-LAMS at the 14th year of mating ( $\approx 0.303\%$  at the 20<sup>th</sup> year of mating) but not significantly ( $p\text{-value} = 0.811$ ). This is better observed in Figure 6.5 (see below) which shows the trend of the TBV mean for both approaches in the last third part of the total period of years of mating.

Taking into account that the larger the size of the population, the bigger the number of males and females available for mating once the equilibrium is reached, i.e. once culling and mortality keep the pool of juvenile and mature candidates about the same amount over the years, the mentioned marginal superiority of LAMS is explained because it is necessary that the number of candidates to select from is enough big to get the expected result for this approach.

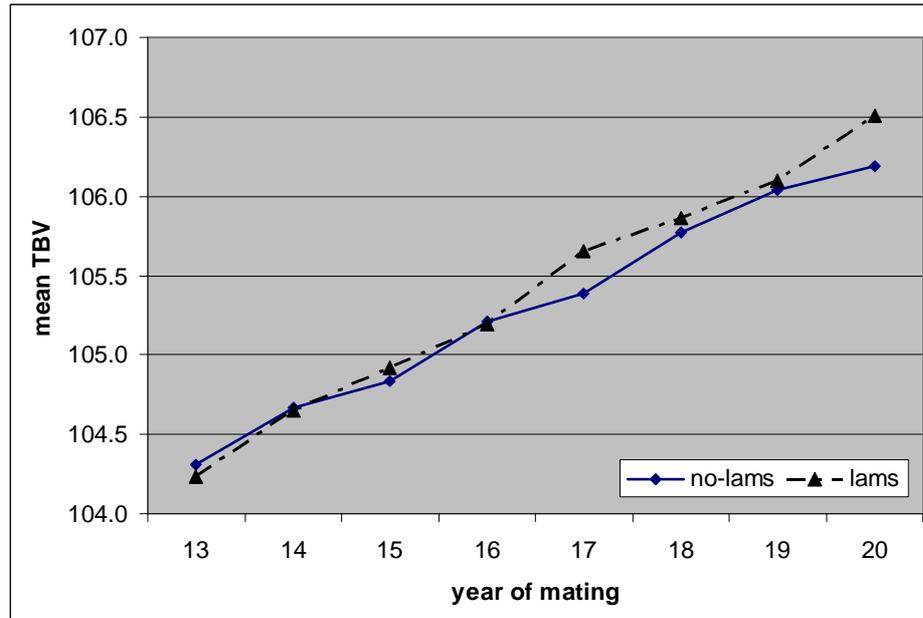


Figure 6.5 Trend of the TBV mean in the last third part of the total period of years of mating for LAMS and no-LAMS

Similar trends in phenotypic response for both approaches appear in Figure 6.6.

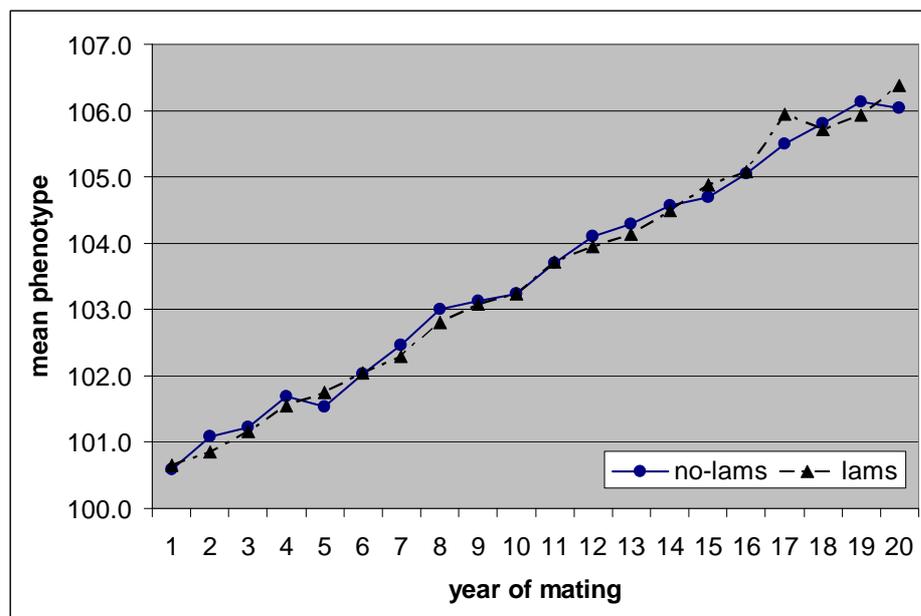


Figure 6.6 Trend of the phenotype mean along the years of mating for LAMS and no-LAMS

Due to the process of selection along the years of mating, 34 selected sires resulted for LAMS and 22 for no-LAMS, while 800 dams were selected for both approaches at the 20<sup>th</sup> year of mating. It is observed in Figures 6.7 and 6.8 (see below) an increment of parental coancestry and the progeny mean of inbreeding respectively over the years of mating. Also it appears in these two figures that these increments are increasingly higher for LAMS than for no-LAMS. Thus, parental coancestry was  $\approx 0.112$  and  $0.091$  for LAMS and no-LAMS respectively at the 20<sup>th</sup> year of mating. While the progeny mean of inbreeding was  $\approx 0.096$  and  $0.077$  for LAMS and no-LAMS respectively at the 20<sup>th</sup> year of mating.

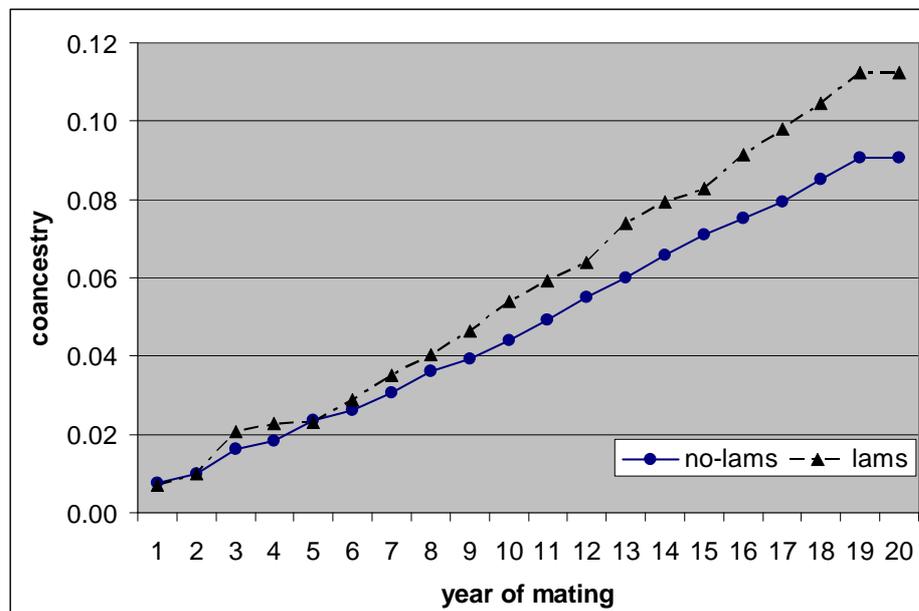


Figure 6.7 Trend of coancestry along the years of mating for LAMS and no-LAMS

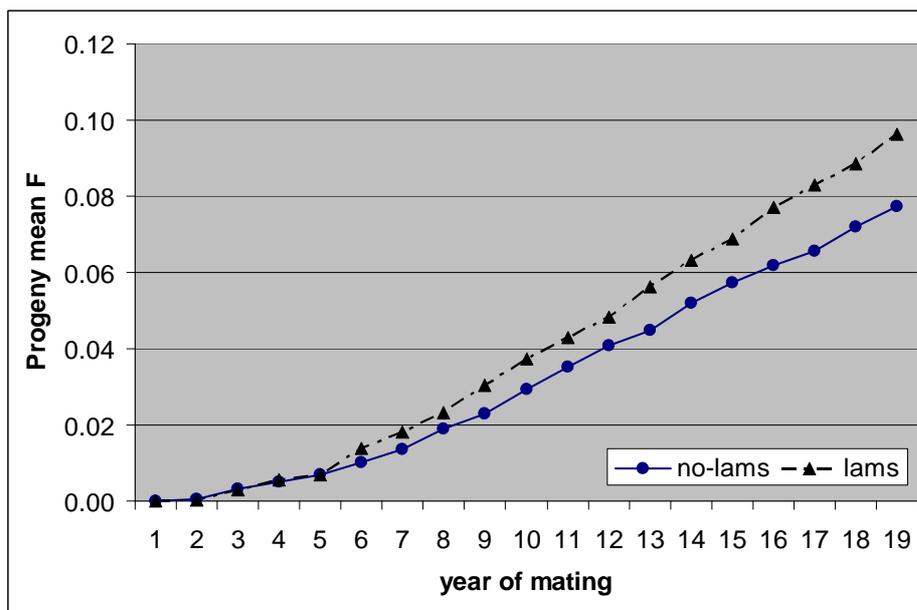


Figure 6.8 Trend of inbreeding along the years of mating for LAMS and no-LAMS

On the other hand, Figure 6.9 shows that the additive genetic variance (VARa) had a very slight decrement for no-LAMS and LAMS over all the years of mating due to the selection process. It is also observed that VARa was higher for LAMS than for no-LAMS.

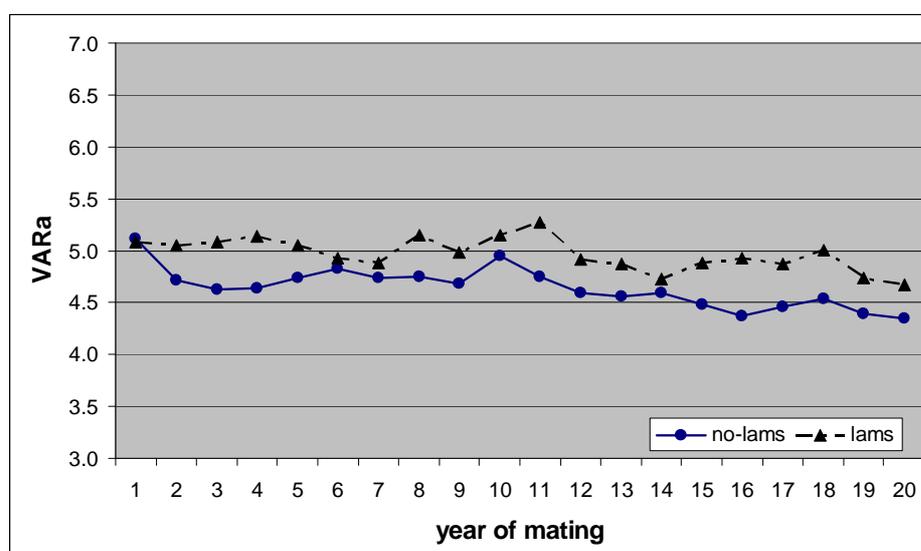


Figure 6.9 Trend of additive variance along the years of mating for LAMS and no-LAMS

Finally, it is observed in Figure 6.10 that the standard deviation of progeny EBVs, i.e. the variance of the progeny merit in the next generation, increased significantly during the first 6-7 years of mating, and became more stable during the final years of mating not only for LAMS but also no-LAMS. Progeny EBVs had a higher standard deviation in LAMS, which was probably due to selecting both more young sires (more genetically variable) and a mix of young and old sires.

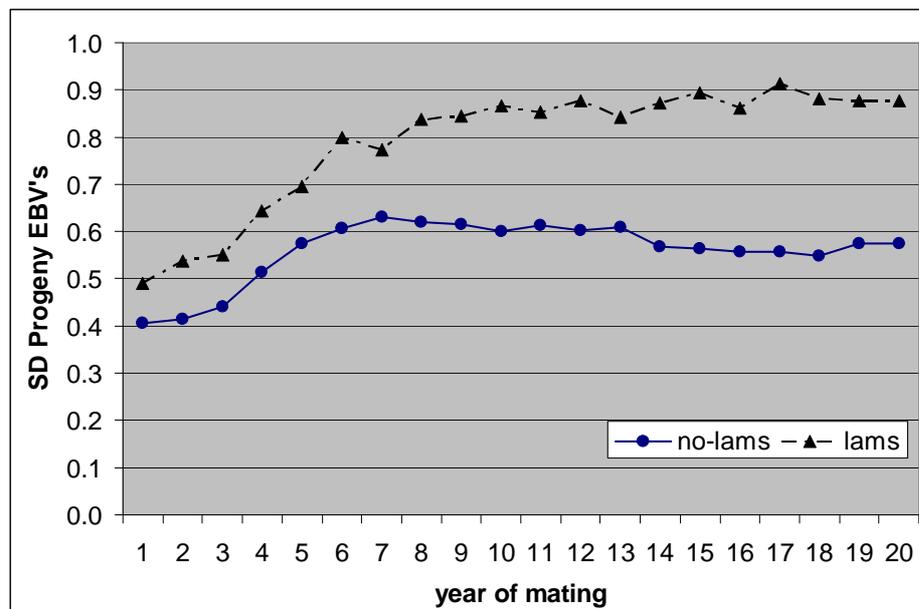


Figure 6.10 Trend of standard deviation of progeny EBVs along the years of mating for LAMS and no-LAMS

The increase in standard deviation of progeny EBVs also suggests more assortative mating, which results in more coancestry/inbreeding. This seems to make sense as mating of more alike parents will give more opportunity for increased selection differential in the next generation, especially for lower selection proportions.

Thus, the higher increment of the variance of the progeny merit in the next generation caused by LAMS allowed bigger selection differential and, in consequence, bigger genetic gains in the grand progeny generation.

Selection of more young sires which are more genetically variable by the tactical approach of LAMS is observed in Figures 6.11, 6.12, 6.13 and 6.14.

In Figure 6.11 it appears that there was a slight trend to select younger sires in LAMS than in no-LAMS along the years of mating.

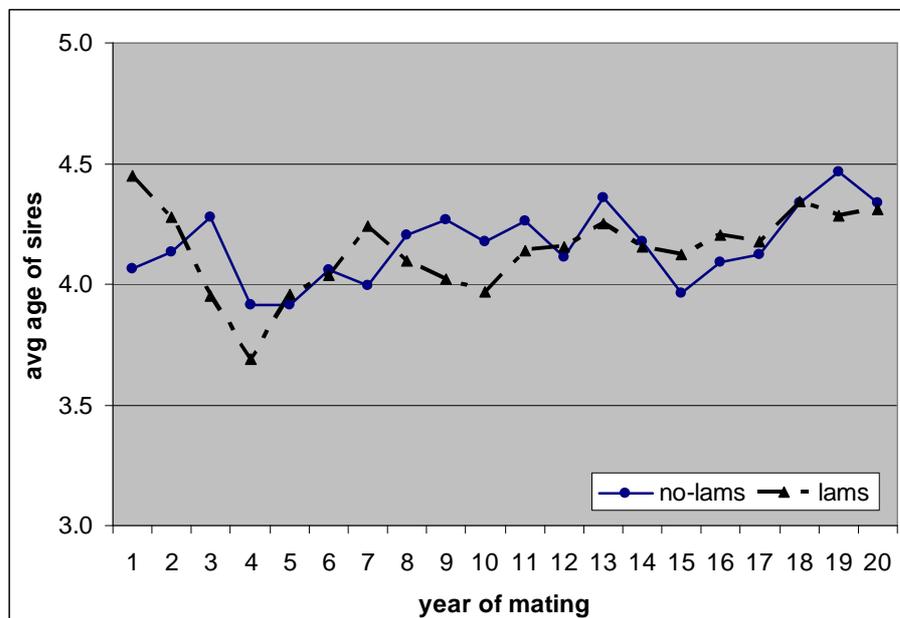


Figure 6.11 Trend of the average age of the selected sires along the years of mating for LAMS and no-LAMS

This is also observed in Figure 6.12, which shows a small higher proportion of 3 and 4 years old selected sires in LAMS (total 63.22%) than in no-LAMS (total 59.34%). The average age of the selected sires was 4.331 in LAMS, while in no-LAMS was 4.344.

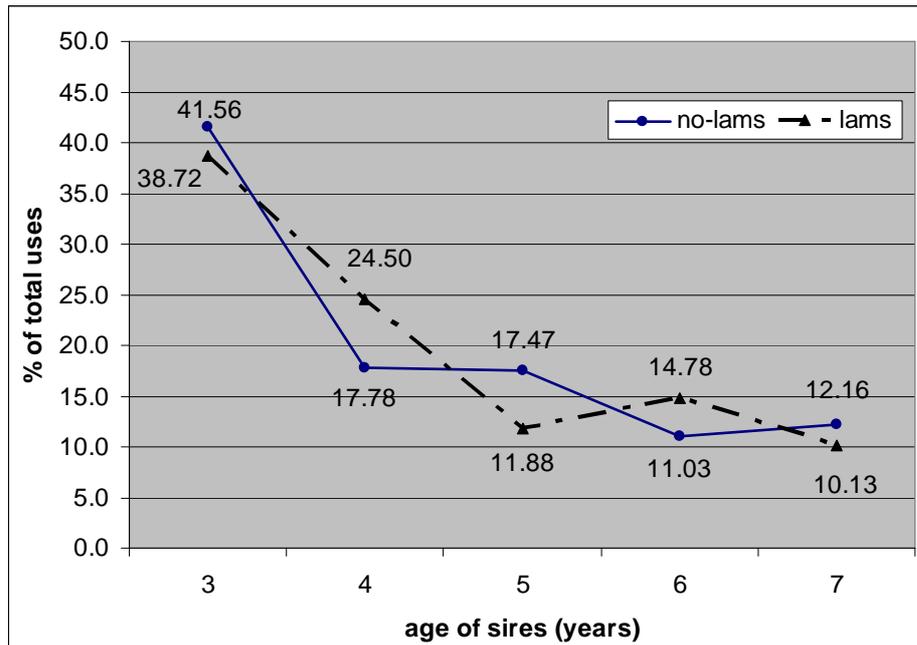


Figure 6.12 Percentage of the total uses by the selected sires according their age for LAMS and no-LAMS

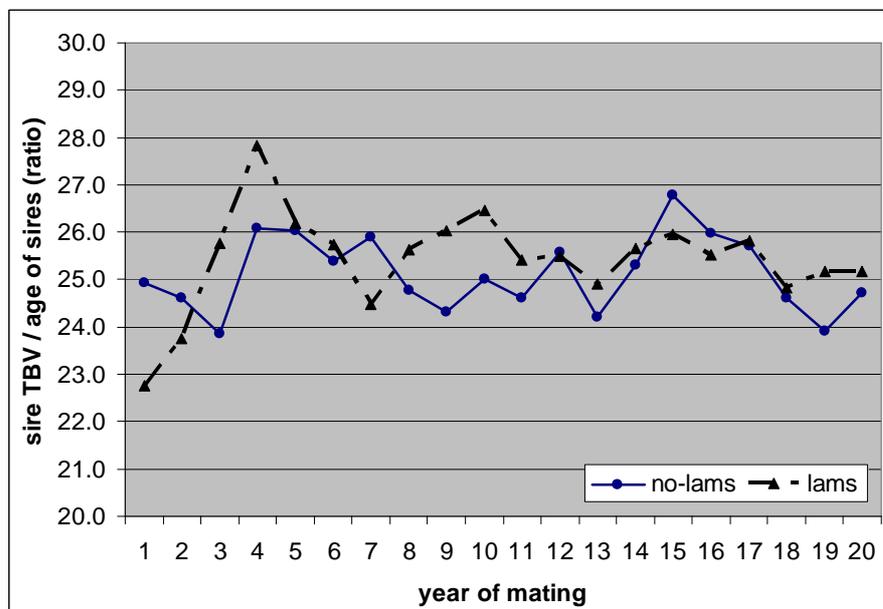


Figure 6.13 Ratio between the average TBV and age of the selected sires along the years of mating for LAMS and no-LAMS

On the other hand, Figure 6.13 shows that the selected sires in LAMS had on average a slightly higher TBV (107.05 units) than those selected in no-LAMS (106.80 units). The observed ratio in LAMS was a consequence of a higher average TBV at a same age of the selected sires or the same average TBV at a lower age of the selected sires. This could explain the slightly higher advantage of the tactical approach of LAMS over no-LAMS.

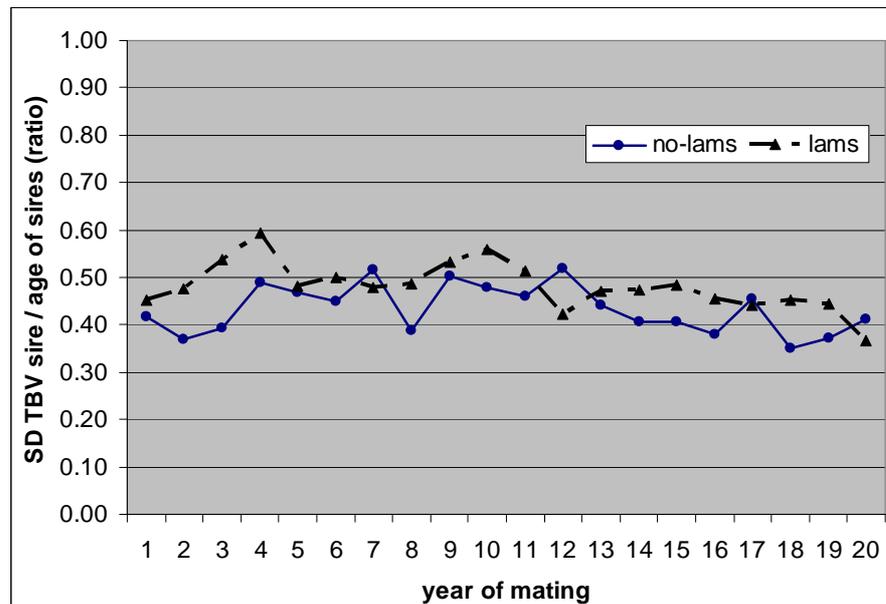


Figure 6.14 Ratio between the average SD TBV and age of the selected sires along the years of mating for LAMS and no-LAMS

The selected sires in LAMS were not only younger but also had a slightly higher average SD of TBV. This is shown in Figure 6.14 in which appears a higher ratio between the average SD TBV and age of the selected sires in LAMS than in no-LAMS. Analogous to Figure 6.13, the observed ratio in LAMS was a consequence of a higher SD of TBV at a same age of the selected sires or the same SD of TBV at a lower age of the selected sires.

From the slightly higher SD of TBV of younger sires selected under LAMS shown in Figure 6.14 it can be deduced that LAMS tends to select younger sires with a lower accuracy and, therefore, with higher variance of the

progeny merit in the next generation, indicating that LAMS favours higher long-term genetic gains in the second generation.

From the above, it could be deduced that the LAMS approach tends towards the setup of a Young Sire Program scenario, but the actual impact on selection and mating patterns seems far from strong.

## **6.6 Discussion**

As mentioned, in the present Chapter, the average increment of the TBV mean per year of mating resulted in 0.289 for LAMS and 0.270 for no-LAMS, i.e. 7.04% higher for LAMS than for no-LAMS. In addition, it appears that the achieved TBV mean starts to be slightly higher in LAMS than in no-LAMS at the 14th year of mating ( $\approx 0.303\%$  at the 20th year of mating) but not significantly (p-value = 0.811).

There seems to be little or no previous work similar to the analysis carried out in this Chapter, probably because of a relative lack of development of mate selection systems that operate beyond control of genetic merit and coancestry.

In our Young Sire Program scenario, investment matings are the use of young sires with lower accuracy than older sires that have progeny. These young sires may also have lower average EBVs than the older sires, with the higher residual genetic variance in their future progeny making them competitive. Realisation matings are then the use of the best young sires in later years, as proven sires with a higher average EBV and higher accuracy. Investment matings are suboptimal in the short term in order to invest for the long term, i.e. to realise the benefits in the latter by revealing residual genetic variation among young sires. Of course, normal selection programs use young sires. Therefore, in fact, the extent to which young sire matings in a YSP program are investment matings, is proportional to the increase in numbers of these matings compared to a normal selection program.

Although the response under LAMS was slightly more favourable in these analyses, the differences involved were far from significant. A YSP is characterised by more selection of younger sires even if they have a lower accuracy. As mentioned, these parameters, despite favourable for LAMS, were not significantly different from those achieved in no-LAMS. The fact that the young sires had a higher TBV than older sires suggests that genetic progress is fast enough to overcome the possibility that they would have lower average EBVs, and be selected mostly because of their collective potential to generate greater gains in the second generation. This latter observation, that young sires had EBVs that were on average higher than older selected sires, indicates that this YSP under the parameters chosen is likely to deviate less from a normal breeding program, where young sires will already be competitive. This in turn gives a reason why the present results for LAMS is little better than for no-LAMS.

Given this, it can be predicted that LAMS will have a bigger impact for programs in which young sires have a lower average breeding value than their older sires that have been subject to more selection. LAMS should have bigger impact on the pattern of use of sires. This would be the case in programs with lower rates of selection response, and in cases where the number of progeny per young sire is rather high, leading to high accuracy in older sires.

The aim of choosing a low heritability for this simulation was to reduce accuracies of animals (actually families). In this way, older males have more distinguished accuracies, which give them the possibility to be selected for longer. So, old and young sires become more differentiated and competitive with each other but for different attributes, i.e. the older males due to their higher accuracy (which allows maximization of the predicted genetic selection differential in the current generation) and the younger males due to their higher prediction error variance (PEV) (which helps to maximize the predicted genetic selection differential in the next generation). This also provides the possible

benefit of using LAMS to optimise use of young sires for traits measured later in life, such as e.g. daughter reproduction, and could be analysed in future studies.

Regarding that the ultimate objective is to do selection in a program that reflects a YSP, this latter effect is favoured with more variation in accuracy (i.e.  $\sigma^2B$ , see above equation [6.3]) between old and young sires. For this, one should look for selecting lowly accurate families of young males through increased  $\sigma^2B$ . This is done essentially through predicted selection among progeny, rather than selection among the same candidates at a later stage in their life.

Following the same idea, the aim of setting the target proportion of selected parents at 40% was to increase the number of parents to select from and favour the possibility of selecting more young parents with lower accuracy but with higher PEV, even though they could have lower EBV than proven sires. Also, that higher setting helps the possibility of a higher variation in accuracy. A lower target proportion, eg. 25%, would increase the selection differential in generation 0 and favour the increment of the average genetic merit in the next generation, i.e. the opposite effect to that which would favour a YSP.

Although proven sires could have on average a higher genetic merit compared to young sires, selection of young sires as candidates increases the residual genetic variance of progeny in the next generation, exploiting the favourable effects of assortative mating. This is due to their lower accuracy of index which implies a higher residual variance of true genetic merit, bringing the possibility to increase the average genetic merit in the second generation. So, through LAMS lower accuracy males are selected that are perceived as competitive because there is chance that the TBVs for some of them are higher than the higher EBV males that they displace. LAMS does this by selecting matings that are predicted to give the biggest selection differential over the current plus next generation.

In Chapter 5, the setup of a YSP was analysed through a deterministic analysis of modelling and simulation. Based on the results for LAMS and no-LAMS in Chapter 5, the percentages of the total usage of the selected sires

according their age are shown in Figure 6.15. Comparing this to the current results in figure 6.12, it can be observed that in both cases LAMS promotes the use of younger sires. However, this is not so clear in the present stochastic analysis, where LAMS has a slightly lower representation in the youngest age class.

It should be noted that in Chapter 5 the age classes refer to the age of mating of males and females while in this Chapter 6 the age of males and females refers to the age of parents when their different progeny were born.

The decrease of the target proportion of selection ( $p$ ) in the analyses of the present Chapter would increase the number of total uses of younger sires at the expense of the number of total uses of older sires.

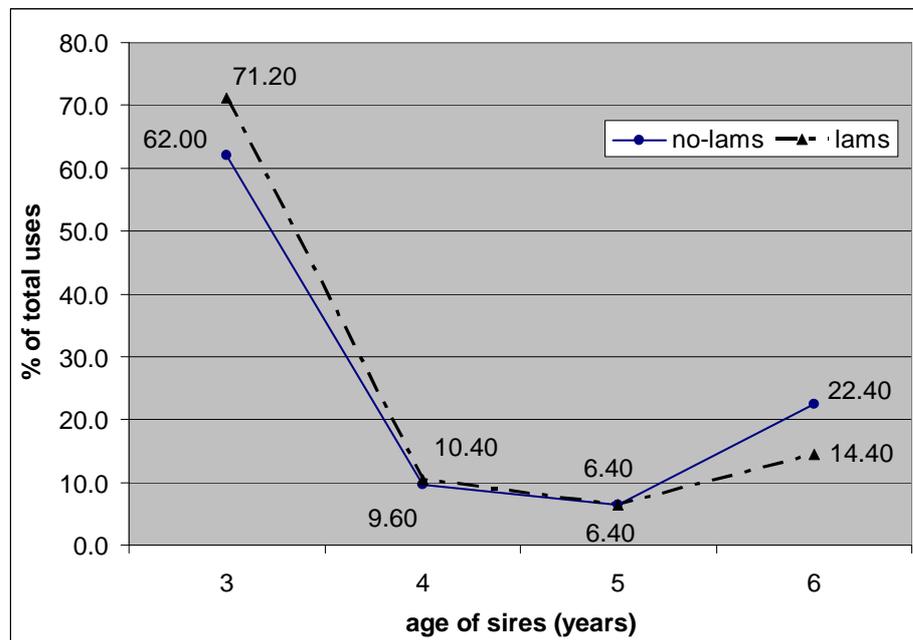


Figure 6.15 Percentage of the total uses by the selected sires according their age for LAMS and no-LAMS using the results of Chapter 5 with 100 proven bulls candidates

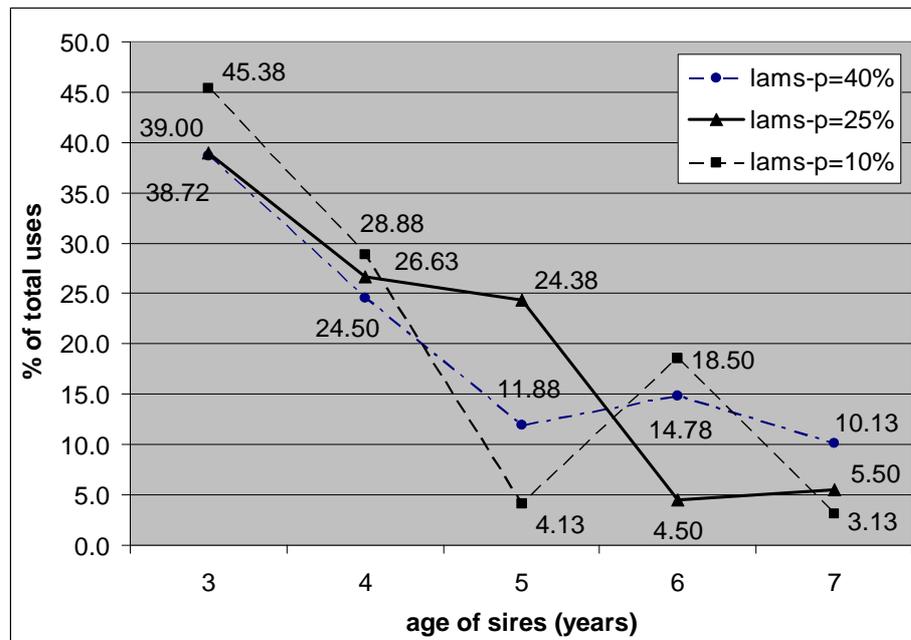


Figure 6.16 Percentage of the total uses by the selected sires according their age for LAMS in Chapter 6 and chosen target proportion of selection ( $p$ )

This is observed in Figure 6.16, which shows, for LAMS, that the proportions of 3 and 4 years old selected sires that are achieved using target proportions of selection equal to 40%, 25% and 10% are 63.22%, 65.63% and 74.25% respectively. This was also noted in the analysis of variation of testing capacity in the case B of Chapter 5. The achieved proportions for 3 and 4 years old selected sires are more robust than for older ages because of larger numbers involved.

In this way, it can be observed how the resulting patterns of the uses of selected sires in this Chapter become closer to the observed pattern of the YSP setup in Chapter 5 as the chosen target proportion of selection is lowered.

Thus, even though there were small differences between LAMS and no-LAMS, tending towards setup of a YSP under LAMS, the impact is very slight compared to imposing a designed YSP directly.

Thus, tactical Mate Selection can in fact set up a scheme that tends towards the features of a YSP. After optimizing using young sires as nucleus sires, a YSP would emerge if that is the best solution (i.e. using many young

sires as nucleus sires). It should be remembered that the objective of Chapters 5 and 6 is to show this. The idea of setting up a YSP is not only due to its advantages, being possibly one of them that using a higher proportion of young sires might be cheaper, but also because it accommodates to the LAMS approach, by which extra genetic gains might be obtained.

Also other differences between the analyses made in Chapter 5 and the present one could be mentioned that could explain these results. Among them could be considered:

- In this Chapter mating distributions or families have been managed instead of cohorts across age classes, like in Chapter 5,
- The deterministic method applied in Chapter 5 does not accommodate deviations from normal distribution curve and there is no random sampling,
- In this analysis, unlike in the one done in Chapter 5, the criterion for YSP operates through predicted selection among progeny, rather than selection among the same candidates at a later stage in their life,
- Inbreeding was accounted for in this analysis, while it was not in the deterministic analysis done in Chapter 5.

Considering that the larger the size of the population, the bigger the number of males and females available for mating once the equilibrium is reached, i.e. once culling and mortality keep the pool of juvenile and mature candidates about the same amount over the years, the mentioned marginal superiority of LAMS is explained because it is needed that the number of candidates to select from to be sufficiently large for this approach to work well. With small herds number of candidates is not sufficiently high, making the LAMS approach non-competitive. This is shown in Figure 6.17, which is equivalent to Figure 6.5, but using 200 dams and 25 sires rather than 800 dams and 100 sires.

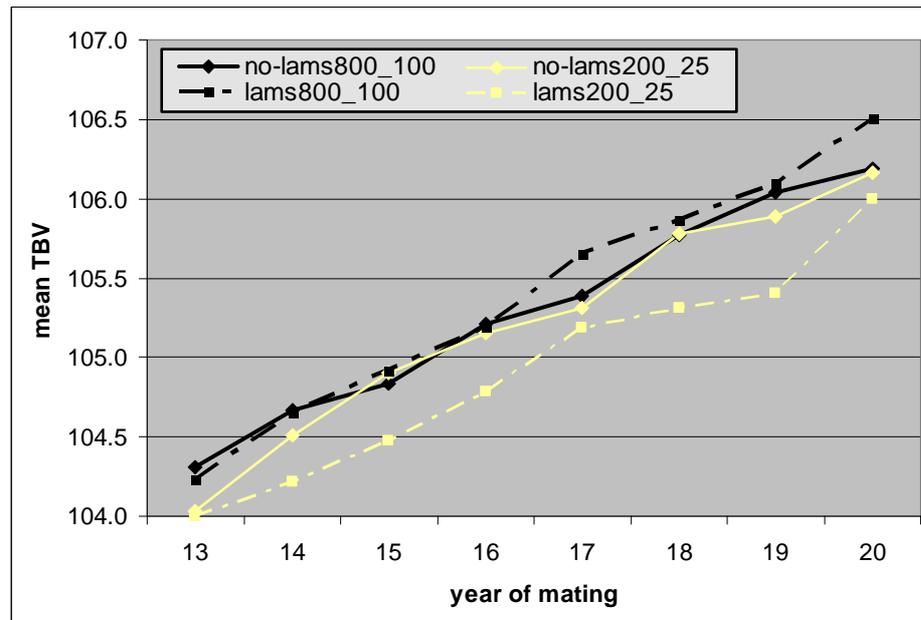


Figure 6.17 Comparison of the trend of the TBV mean in the last third part of the total period of years of mating for LAMS and no-LAMS between using an initial herd with 800 dams and 100 sires or other with 200 dams and 25 sires

Also, as it is already known, an increase in herd size causes a decrease in inbreeding and coancestry and allows an increase in the variance of the progeny merit in the next generation, which, in the case of LAMS, favours higher long-term genetic gains in the second generation.

In this way, it appears that higher genetic gains in LAMS are favoured in herds of higher size. This was also observed in the analysis done in Chapter 5 (see section 5.5.1 case C).

Even with infinite population size more gain is expected from LAMS. This is because it exploits residual genetic variation that can be used for selection at later ages, rather than circumvent inbreeding/coancestry.

However, these comments should not be taken to imply that a sufficiently large herd size would necessarily make LAMS usefully better than no-LAMS.

Mean levels of inbreeding and coancestry at the 20<sup>th</sup> year of mating were relatively low due to the large size of the herd used in this analysis, but levels were higher for LAMS. So, even though the number of selected sires at the 20<sup>th</sup>

year of mating was substantially higher for LAMS (34 selected sires) than for no-LAMS (22 selected sires), the higher mean inbreeding and coancestry achieved by LAMS implies that the selected parents for this approach are significantly more genetically related.

There are several YSP that are being carried out by different companies and breeder associations, mainly in Australia and U.S.A. Two of the YSP that are carried out in Australia are based on progeny testing and report their rate of genetic gain estimates. These are Meat Elite Australia (MEA) which is for meat sheep and the one carried out by Angus Australia which started in 2004. MEA reports that it is achieving 21% faster genetic progress than the average for all terminal sire groups across Australia for a range of profit-driving traits, although this benefit probably involves other aspects of the breeding program, such as accurate recording. Angus Australia has the objective of doubling the rate of genetic gain within the following 10 years, i.e. before or in 2014.

A fair comparison between the results in this Chapter and the estimations of the rate of genetic gain in the two mentioned YSP carried out in Australia would not be feasible as in MEA the reported percentage genetic gain comes from the difference between a subpopulation and the rest of that population and, in the case of Angus Australia, the stated objective does not appear to be based on a prediction of the impact of the YSP per se.

## **CHAPTER 7**

### **GENERAL DISCUSSION**

## ***7.1. Summary***

Frequently, optimisation of breeding program design may involve complex problem solving. Through strategic optimisation the whole population model of the breeding program accommodates parameters from which emerge strategies to determine what traits to be measured, what are the economic weights associated with them, and rules to help manage genetic gain, inbreeding, use of reproductive techniques, etc.

Aiming to the same purpose, alternative tactical optimisation approaches have been developed and implemented, which determine the pattern of use of prevailing animal resources. Through these tactical methods a function is maximized that involves the genetic response and other components to be implemented through the selection and assignment of matings and using genetic algorithms.

It would be appropriate to note that the implementation “costs” of strategic and tactical optimisation approaches are different: in the strategic approach, the rules or parameters are determined at intervals or only once, whereas for tactical, it is essentially a continuous process. Any benefits of the tactical approach, in genetic gain over some period, must be set against any additional costs incurred over that period. It can also be pointed out that once a tactical approach has been invested in, there are many extra technical, logistical and cost issues that can be added to the objective function. In practice, benefit should accrue from a number of directions.

This thesis presents the comparison of results of applying tactical and strategic approaches for different breeding programs according to diverse objectives targeting commercial beef production systems in Argentina, with the use of genetic algorithms for Mate Selection in a pure beef cattle population. Among these genetic algorithms notably Differential Evolution has been illustrated to optimise breeding program design, breeding objectives, and as a selection tool.

In Chapter 3, through the analysis of some simple examples of Mate Selection based on a simulated population of beef data, it is shown that optimal balance between high genetic gain and low parental coancestry, when selecting breeding animals, can be achieved with the optimal set of contributions of the selected individuals' genes to the breeding program. The consequences of the management of progeny inbreeding and variance of progeny, plus other factors, showing the impact of different policies, constraints and parameter settings are also tested. A simple case of Look-Ahead Mate Selection (LAMS) is also included as a preliminary platform for investigating the setting up of a Young Sire Program (YSP), using this approach in Chapters 5 and 6. As a result, it is shown that there is a marginal positive result in genetic response when it is aimed at the grand progeny generation by maximizing the predicted genetic selection differential to be achieved in the progeny generation (i.e. LAMS) (following Kinghorn and Shepherd, 1994 and Hayes et al., 1998) compared to aiming to improve the immediate next generation (no-LAMS).

A YSP involves mating young sires to generate test progeny to accurately test them for selection at older ages in a breeding program, which, in turn, allows faster genetic progress in the herd. On the other hand, herd size affects genetic response due to its relation to selection intensity and inbreeding. Joining a cooperative YSP appears as a suitable solution for stud-farms with a limited herd size in Argentina to improve the rate of genetic gain without increasing inbreeding.

Chapter 4 investigates the consequences of applying a tactical mate selection approach in stud-farms with a moderate to large herd size that apply a two-tier nucleus breeding scheme, using real data and considering various costs and constraints. Results for both cases, the use of diverse reproductive techniques and dissemination of breeding males to different beef markets, illustrate that the application of Mate Selection is a tactical approach suitable to simultaneously select different sets of sires in the stud-herds according to the particular breeding objectives of different commercial units, even including

those of the stud. The consideration of these objectives in the Mate Selection Index function allows the optimisation of profit over herds, in addition to the control of other genetic aspects such as maintenance of genetic diversity and avoidance of inbreeding.

In Chapter 5 different aspects of the implementation of this YSP are analysed through a deterministic modelling and simulation. The LAMS approach results in a marginally higher, but not significant, genetic response along years compared to aiming to improve the immediate next generation (no-LAMS). This marginally higher genetic gain increases as the number of sire candidates over age classes increases. Around 3.0% more long-term genetic gain is obtained in LAMS than in no-LAMS when 100 proven bulls are available as candidates to be selected as sires. The distribution of elite sires over age classes is also affected, with a higher selection of young candidates as sires.

In Chapter 6 the consequences of applying a LAMS approach in this YSP scenario are tested by stochastic simulation, trying to show whether tactical Mate Selection can in fact set up such a scheme, and whether it improves longer-term response by making opportunistic use of the prevailing animal resources. The resulting average selection response in TBV was 0.289 for LAMS and 0.270 for no-LAMS (i.e. 7.04% higher). Also it appears that the achieved TBV mean starts to be slightly higher in LAMS than in no-LAMS at the 14th year of mating ( $\approx 0.303\%$  at the 20th year of mating) but not significantly ( $p\text{-value} = 0.811$ ).

Comparing the resulting patterns of the uses of selected sires from Chapter 6 to those in Chapter 5, it can be observed that these patterns in Chapter 6 become closer to the observed ones in Chapter 5 as the chosen target proportion of selection is decreased. Despite the small differences between LAMS and no-LAMS, tending towards setup of a YSP under LAMS, the impact is very slight compared to imposing a designed YSP directly. Thus, tactical Mate Selection can in fact set up a scheme that tends towards the features of a YSP.

## ***7.2. Discussion***

The techniques used in this thesis for investigation are based on methods for progressive implementation of breeding programs, and as such bear direct relevance to progressive sectors of the beef industry.

There seems to be little or no previous work similar to that carried out in Chapters 3, 4, 5 and 6, probably because of a relative lack of development of mate selection systems that operate beyond control of genetic merit and coancestry.

The analyses of the various cases shown throughout this thesis illustrate that the application of Mate Selection is a dynamic tactical approach suited to optimise breeding program designs when more complex objectives are desired by breeders. Thus, Mate Selection can have relevant impact for some issues, including budgetary and logistical constraints, that are currently difficult to handle simultaneously (multiple objectives, the use of different reproduction techniques, control of inbreeding, maintenance of genetic diversity to target multiple beef markets, etc., being possibly all these even different for nucleus than for other tiers of the herd).

In order to accommodate antagonistic objectives it is necessary to make trade-off decisions, simultaneously balancing the science applied and the experience and preferences of the breeder. In this way “satisfaction” becomes a very important aspect in the Mate Selection process and the final achieved solution will be surely implemented and will not only be taken just as a recommendation (Kinghorn et al., 2002; Gondro and Kinghorn, 2007; Kinghorn et al., 2009). So, the interaction between the practitioner with the optimisation at runtime and with the breeder in this process gives the opportunity not only to achieve an optimal solution but also a solution that “satisfies” the breeder’s desires, with some ‘ownership’ of the solution.

It should be noted that while breeders may know their intended selection objectives, involving trait changes to increase economic efficiency, it is necessary to set a whole population model of the breeding program. Through strategic optimisation the model accommodates parameters from which strategies emerge to determine what traits to be measured, what are the economic weights associated with them, and rules to help manage genetic gain, inbreeding, use of reproductive techniques, etc.

Frequently, optimisation of breeding program design may involve complex problem solving such as maximizing response while maintaining under control the costs of the program associated with the use of each breeding animal, the use of reproductive technologies, and the rate of inbreeding. Some flexible methodologies have been developed in this field considering optimal mating structures for all animals of the population, and all are implemented through sets of rules to follow. Alternatively, with the same purpose, tactical optimisation approaches, which determine the pattern of use of prevailing animals, have been developed and are being implemented in the industry. The latter can use inputs such as economic indicators obtained with EBVs, the pedigree and reproduction costs, maximization of a function that involves the genetic response and other components to be implemented through the selection and assignment of matings. The choice of which traits to measure or to genotype on which animals is the only key issue not covered by the combined Mate Selection approach. This has to be handled by extension to the method.

It would be appropriate to note that the implementation “costs” of strategic and tactical optimisation approaches are different: in the strategic approach, the rules or parameters are determined at intervals or only once, whereas for tactical, it is essentially a continuous process. Any benefits of the tactical approach, in genetic gain over some period, must be set against any additional costs incurred over that period. It can also be pointed out that once a tactical approach has been invested in, there are many extra technical, logistical

and cost issues that can be added to the objective function. In practice, benefit should accrue from a number of directions.

Tactical methods obtain solutions to any of these problems with the use of genetic algorithms.

Genetic algorithms have been successfully applied to breeding program design (as shown by Kinghorn, 1998; Meszaros et al., 1999; Kinghorn et al., 2002).

Throughout this thesis the use of genetic algorithms has been illustrated, notably Differential Evolution, to optimise breeding program design, breeding objectives, and as a selection tool.

In Chapter 3 the results of some simple examples of Mate Selection were presented using the Pedigree Viewer system (<http://www-personal.une.edu.au/~bkinghor/pedigree.htm>) based on a simulated population of beef data. This program finds the optimal solution by using a stochastic search algorithm based on the Differential Evolution (DE) algorithm. Additional attention to mate allocation, using simultaneous Mate Selection, adds management of inbreeding and genetic variance.

Through the analysis of the given examples, it was shown that optimal balance between high genetic gain and low parental coancestry, when selecting breeding animals, can be achieved with the optimal set of contributions of the selected individuals' genes to the breeding program. The consequences of the management of progeny inbreeding, progeny variance and predicted grand progeny merit, plus other factors were also tested, showing the impact of different policies, constraints and parameter settings using the same simulated beef data.

In this way, after the analyses, one was better prepared to understand consequences in a herd when optimisation of the breeding program design is looked for using more complex objective functions.

In addition to the mentioned examples, a simple case of Look-Ahead Mate Selection (LAMS) was included as a preliminary platform for

investigating the setting up of a Young Sire Program using this approach in Chapters 5 and 6.

As a result, it was shown that there was a slight marginal positive result in genetic response when the grand progeny generation is targeted by maximizing the predicted genetic selection differential to be achieved in the progeny generation (i.e. the look-ahead scheme) (following Kinghorn and Shepherd, 1994 and Hayes et al., 1998) compared to aiming to improve the immediate next generation, which is the traditional scheme applied (no-LAMS).

In Chapter 4 the consequences of applying a tactical mate selection approach were observed in stud-farms with a moderate to large herd size that apply a two-tier nucleus breeding scheme, using a real data base from advanced Angus purebred stud-farms of Argentina, considering various costs and constraints.

Results for both cases, showed how the use of diverse reproductive techniques considering budgetary constraints and dissemination of breeding males to different beef markets can be optimised. In the latter case it is shown that the application of Mate Selection is a tactical approach suited to simultaneously select different sets of sires in the stud-herds according the particular breeding objectives of different commercial units, even including those of the stud. The consideration of these objectives in the Mate Selection Index (MSI) function allows the optimisation of profit over herds, in addition to the control of other genetic aspects such as maintenance of genetic diversity and avoidance of the undesirable consequences of inbreeding.

It should be taken into account that all the cases in Chapters 3 and 4 were analysed after only one round of selection. As mentioned, no-LAMS and LAMS approaches were applied in a Young Sire Program scenario over multiple selection rounds in Chapters 5 and 6. Also, it should be noted that the LAMS approaches used in Chapter 3, on one hand, and Chapters 5 and 6, on the other, differ. In this way, in Chapter 3, LAMS aimed for the maximization of the selection differential in the grand progeny generation, which was predicted by

the difference between the average merit of their selected parents and the average merit of the entire population of current candidates before they were selected. To do this, the progeny predicted to be born from the current selected candidates were themselves put through a selection process, and the means of these selected progeny used to predict mean grand progeny merit. Chapters 5 and 6 also aimed for the maximization of the grand progeny mean, but the latter was predicted with use of the prediction error variance of EBVs plus the average inbreeding coefficients (these latter being only considered in Chapter 6) of the selected parents, to exploit higher residual genetic variance, as deviated from parent mean EBV, in progeny families with lower accuracy parents. As already mentioned, in Chapters 5 and 6 simulations were done over multiple selection rounds, being the approach fully tactical as the solution was optimised at every selection round.

A Young Sire Program (YSP) involves mating young sires to generate test progeny outside the nucleus. One of the main objectives of a progeny test is to accurately test the young sires for selection at older ages in a breeding program. After optimizing using them as nucleus sires, a YSP would emerge if that was the best solution (i.e. using many young sires as nucleus sires). Progeny testing could be an expensive exercise for individual studs. Thus, joining a cooperative YSP appears as a suitable solution for stud-farms with a limited herd size in Argentina, decreasing the cost of genetic progress compared with the use of imported semen. Moreover, commercial clients benefit from more accurate breeding values and better genetics.

In Chapter 5 investigates whether LAMS could give more genetic gain than truncation selection in the longer term and to check whether in an optimal breeding program the proportion of young sires to be selected should be increased compared to common practices.

Thus, in Chapter 5 different aspects of the implementation of a YSP were analysed through a deterministic modelling and simulation. Optimal solutions for various scenarios in terms of proportion of young bulls selected for the

nucleus and number of young bulls mated for the purpose of progeny testing were shown in this analysis. One of the main objectives of a progeny test is to accurately test the young sires to use in a breeding program.

Setting a whole population model of the breeding program, breeding alternatives were compared from which different strategies were derived.

Although the number of age classes for males and females set in the model were the same, the approach used to modelling age structure is flexible since that number can be changed, even using different number of age classes by sex, and calculating the different parameters accordingly.

The use of a DE algorithm was the key tool to determine the optimal proportions of proven and young bulls to be selected as parents along age classes and to balance genetic parameters. The application of truncation selection based on BLUP EBV to optimise the selection of sires in YSP, i.e. to reserve the optimal proportion of matings for young and proven sires, is a strategic optimisation, involving simply selecting the sires with the best breeding value (i.e. no-LAMS). Meanwhile, the LAMS selection method looks for selecting relatively more young sires to have progeny due to the variation among sires' true breeding values. This variation would increase when selecting more young sires as young sires are selected with less accurate estimated breeding value but with more prediction error variances. This is achieved by maximising the predicted genetic merit in the grand progeny simultaneously considering merit both one and two generations ahead through the use of a DE algorithm. Thus, LAMS appears as a tactical optimisation given that more variation would result in more selection differential and more genetic gain in the long term.

The difference of accuracy between older and younger bull age classes and its impact on the distribution of estimated breeding values as well as the mean of selected sires appears to have a major effect when selecting across age classes, and consequently on the achieved genetic response.

Both selection differential and PEV of each age class are determined by the EBV accuracy of the animals of that age class. Under no-LAMS, a higher accuracy in an age class corresponds to a higher selection differential among candidates (at equal selection intensities) and, therefore, a larger number and proportion of selected animals from that age class. Under LAMS, a higher accuracy in an age class corresponds to a lower PEV (at equal genetic variances), and this promotes a smaller number and proportion of selected animals predicted to be selected among progeny, due to the lower residual genetic variation about predicted family means in that age class. So, the higher the difference in accuracy between age classes, the more the no-LAMS approach will tend to increase the number and proportion of selected animals in the age class with higher accuracy, while LAMS approach will tend to increase the number and proportion of selected animals in the age class with lower accuracy.

As mentioned, one of the main objectives of a progeny test is to accurately test the young sires to use in a breeding program. The achievement of this goal allows, in turn, higher genetic gains along selection rounds, which is one of the main objectives of a breeding program. All this was shown through results where, in a fixed sized nucleus of 1000 combined females brought by 10 small studs participating in this YSP, this aim is achieved by increasing the proportion of selected young bulls as sires rather than increasing the number of young sires as a result of increasing the total amount of selected sires. When LAMS is compared to no-LAMS, the results in LAMS showed a larger proportion of selected sires from the young bull age classes and a slightly higher rate of genetic gain and, therefore, slightly higher means after 50 rounds of selection. This marginally higher genetic gain increases as the number of progeny tested 2-year-old bulls increases which, in turn, causes the increase in the number of bulls in the following older age classes as candidates. This can be achieved by an increase in the test capacity outside the nucleus. Thus, a genetic response equal to 0.200 TBV genetic standard deviations per year was obtained in the traditional approach (no-LAMS) and around 3.0% more genetic gain was

obtained applying a look-ahead mate selection scheme approach (LAMS) compared to no-LAMS when 100 proven bulls are available as candidates to be selected as sires. Although the mentioned differences between LAMS and no-LAMS become even larger as larger is the number of proven sires to have available as candidates, with LAMS giving around 5.9% more annual gain with 200 proven sires and around 8.6% more annual gain with 450 proven sires, it is necessary to take into account that the resources and costs to obtain larger numbers of proven sires would increase considerably.

So, results showed that there was a marginal positive result in genetic response when the grand progeny generation is targeted by maximizing the genetic selection differential to be achieved in the progeny generation (LAMS) compared to aiming to targeting the immediate next generation (no-LAMS).

Inbreeding was not taken into account in the deterministic analysis of Chapter 5 and could be considered as a shortcoming because inbreeding reduces the genetic variance, causing an overestimation of the achieved genetic gains. Estimation of inbreeding rate and its inclusion in optimisation would be important if the model was used to optimise a whole breeding population where inbreeding could become a problem. Thus, the calculation of inbreeding and its effects on all genetic parameters and the corresponding consequences in the results of YSP should certainly be considered in small nuclei and/or nuclei with relatively few total selected sires. However, considering the size of the herd and the number of selected sires used in Chapter 5, inbreeding appears to be low. It should be noted that inbreeding can be controlled by optimal contribution selection which was applied in the stochastic analysis done in Chapter 6.

Although the size of the nucleus has been varied (as well as the number of sires and females included in it) in the cases analysed in Chapter 5, optimisation of the size of the nucleus would be a useful and straight forward extension to the model, however due to the increased potential for inbreeding with a small nucleus it would be advisable that constraints to inbreeding rate were also included if this was to be considered.

Taking into account that, in practice, the Bulmer effect will impact on response to selection in the medium term by reducing genetic variance, this effect was considered in the analyses done in Chapters 5 and 6. Given that most of the traits in the breeding objective would be under selection for more than 20 years, it is highly likely that the variance in the traits under selection will be less than the base genetic parameters from the animal model used to estimate them, thus the genetic merit predicted would be higher than that considering the Bulmer effect. Thus, results in Chapter 5 showed a reduction in variance due to selection mainly during the first 10 mating seasons, becoming more stable after that point. Also work in Chapter 5 showed that the predicted rate of genetic gain would be 25.9% overestimated high if the Bulmer effect had been overlooked. As the calculated genetic means affect selection relatively more than accuracies, this causes more young sires to be selected which, in turn, causes a lower average generation interval. It could be added that, in reality, earlier breeding values are more determined by family information, and it is the between family variation that is reduced by the Bulmer effect. As this was a deterministic simulation, correlation of EBVs among relatives was not accounted for. It would have needed a stochastic simulation to account for both this and inbreeding. That correlation was also accounted for by use of stochastic simulation in the analysis of Chapter 6.

Finally, in Chapter 6 the consequences of applying a look-ahead mate selection scheme were tested in this YSP scenario, as mentioned, by stochastic simulation, trying to show whether tactical Mate Selection can in fact set up such a scheme, and whether it improves longer-term response by making opportunistic use of the prevailing animal resources.

For this, in the process of Mate Selection some weight is implicitly given to investment matings, which compete with realisation matings when maximizing predicted grand progeny merit. All this may reduce selection differential in the parent generation, as a consequence of selection of animals which would not be selected on EBV alone but can contribute sufficiently to a

mating design which takes advantage of that better information. In this way, variance in predicted genetic merit in the progeny generation is increased and this will generally be reflected by increased variance in true genetic merit. Even though a reduction in one-generation response is likely, this results in an increase in two-generation response, i.e. an increase in genetic merit in the second generation.

Due to selection, the resulting average response in TBV was 0.289 for LAMS and 0.270 for no-LAMS, i.e. 7.04 % higher for LAMS than for no-LAMS. Also it appears that the achieved TBV mean starts to be slightly higher in LAMS than in no-LAMS at the 14th year of mating ( $\approx 0.303\%$  at the 20th year of mating) but not significantly ( $p\text{-value} = 0.811$ ).

The marginal superiority of LAMS appears only when the number of candidates to select from is sufficiently high in the herd. Otherwise, LAMS is not competitive. Also, as it is already known, the increase in the size of the herd causes a decrease in inbreeding and coancestry and allows an increase in the variance of the progeny merit in the next generation, which, in the case of LAMS, favours higher long-term genetic gains in the second generation. In this way, it appears that higher genetic gains in LAMS are favoured in herds of higher size. This was also observed in the analysis done in Chapter 5 (see section 5.5.1 case C). However, these comments should not be taken to imply that a sufficiently large herd size would necessarily make LAMS usefully better than no-LAMS.

Mean levels of inbreeding and coancestry at the 20th year of mating were relatively low due to the large size of the herd used in the stochastic analysis, but levels were higher for LAMS. It should be noted that the size of the herds used in Chapters 5 and 6 were similar.

A YSP is characterised by more selection of younger sires, even if they have a lower accuracy. The fact that the young sires actually had a higher TBV than older sires suggests that genetic progress is fast enough to overcome the possibility that they would have lower average EBVs, and be selected mostly

because of the potential of these young sires to generate greater gains in the second generation. This latter observation, that young sires had EBVs that were on average higher than older selected sires, indicates that this YSP under the parameters chosen is likely to deviate less from a normal breeding program, where young sires will already be competitive. This in turn gives a reason why the present results for LAMS is little better than for no-LAMS.

Given this, it can be predicted that LAMS will have a bigger impact for programs in which young sires have a lower average breeding value than their older sires that have been subject to more selection. LAMS should then have bigger impact on the pattern of use of sires. This would be the case in programs with lower rates of selection response, and in cases where the number of progeny per young sire is rather high, leading to high accuracy in older sires.

Choosing of a low heritability for this simulation to make old and young sires become more differentiated and competitive with each other in order to help to maximize the predicted genetic selection differential in the next generation, provides the scenario for a possible benefit of using LAMS to optimise use of young sires for traits measured later in life, such as e.g. daughter reproduction, which could be analysed in future studies.

On the other hand, comparing the resulting patterns of the uses of selected sires in Chapter 6 to those in Chapter 5, it can be observed that these patterns in Chapter 6 become closer to the observed ones in Chapter 5 as the chosen target proportion of selection is decreased.

Even though there were small differences between LAMS and no-LAMS, tending towards setup of a YSP under LAMS, the impact is very slight compared to imposing a designed YSP directly. Thus, tactical Mate Selection can in fact set up a scheme that tends towards the features of a YSP.

Although percentage increases in gain were small in the analyses done in the experimental Chapters 3, 4, 5 and 6, accumulated benefit over a longer horizon could be sufficient to make implementation worthwhile. A cost-benefit analysis of implementation has not been attempted in this thesis.

This thesis has taken an illustrative approach to the testing the value of Mate Selection for exploiting a range of opportunities in genetic improvement. It has been difficult to be fully evaluative, due to the high cost of replication in such involved analyses, difficulty in defining appropriate control treatments, and because of difficulty in fully reflecting reality.

Value for most issues seems to have been illustrated reasonably well, with the exception of LAMS as interpreted and set up here, where benefits are marginal, even under conditions favourable to LAMS.

Proper evaluation of the benefits likely to accrue in practice is most simply done at the time of real implementation. This can be done by running Mate Selection analyses on prevailing real data sets, and comparing predicted results (progeny merit, parental coancestry, costs, etc.) to those predicted under the use of other classical or 'control' implementations.

The use of genetic algorithms through suitable ad hoc software to optimise breeding program design and, in general, to solve complex problems in animal breeding, appears promising for the progressive breeders of Argentina, and it is expected that it becomes a widely applied dynamic tactical tool in the beef industry.

Different companies and breeder associations in Australia, U.S.A. and other countries announce they are carrying out YSP, not only for beef cattle but also for dairy cattle and other species. All of them are based on progeny test in order to select the best young sires. However, it is not clear if they refer to the concept of YSP as it was done in this thesis. Notable schemes are Meat Elite Australia (MEA) for meat sheep (which is carried out by an Australian sheep breeding group) and the YSP carried out by Angus Australia which started in 2004. MEA reports that it is achieving 21% faster genetic progress than the average for all terminal sire groups across Australia for a range of profit-driving traits, although this benefit probably involves other aspects of the breeding program, such as accurate recording. On the other hand, Angus Australia has the objective of doubling the rate of genetic gain within the following 10 years, i.e.

before or in 2014. Apart from these two groups, the others do not report their rate of genetic gain estimates.

It would not be possible to rationally compare the estimates of the rate of genetic gain in these two YSP to the results in Chapters 5 and 6 of this thesis. The reason for this is that in MEA the reported percentage genetic gain comes from the difference between a subpopulation and the rest of that population and, in the case of Angus Australia, because their predicted rate of genetic gain appears as an objective and not as a prediction of the impact of the YSP that is being carried out.

## REFERENCES

Allaire F.R., 1993. Intra-breed mating pair selection methods for improvement of nonlinear merit from additive genetic inheritance. *Journal of Dairy Science* Vol. 76, No. 8, 1993

Amer, P.R., Fox, G.C. and Smith, C., 1994. Economic weights from profit equations: appraising their accuracy in the long run, *Animal Production*, 58:11-18.

Avendaño, S., Villanueva, B. and Woolliams, J., 2003. Expected increases in genetic merit from using optimized contributions in two livestock populations of beef cattle and sheep. *Journal of Animal Science*, 81:2964–2975.

Ávila, H., 2006. Consideraciones sobre aspectos relacionados a la búsqueda de “Calidad Industrial” y mejoras al actual sistema de Tipificación. 2º Curso de Capacitación y Certificación de Ecografistas para la Toma de Datos de Rasgos de Calidad de Carne. Asociación Argentina de Angus. Buenos Aires, Argentina.

Ávila H., Lagos F. and Lasta J., 2004. Terminación Uniforme de Tropas y su Relación con la Calidad de la Carne. Conference. Marca Líquida Agropecuaria, Córdoba, Argentina.

Bäck, T., Ed. 2003. *Handbook of Evolutionary Computation*. Institute of Physics Publishing, Bristol.

Banks, R.G., van der Werf, J.H.J. and Kinghorn, B.P., 1998. Optimal sire use policies in sheep breeding. In Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, Armidale. Vol. 24: 15-18.

Barwick, S., 1993. B-OBJECT: A PC-Program to derive economic weights for beef cattle. In Fewson, D., James, J., Nitter, G., Kinghorn, B., Barwick, S.,

Grasser, H.-U. and Savicky, J., editors, Design of livestock breeding programs, pages 79–92. Animal Genetics and Breeding Unit. The University of New England. Armidale. Australia.

Berg, P., Nielsen, J. and Sørensen, M.K., 2006. EVA: Realized and predicted optimal genetic contributions. In Proceedings of the 8th World Congress on Genetics Applied to Livestock Production [CD-ROM communication 27-09].

Bijma, P., 2000. Long-term genetic contributions: Prediction of rates of inbreeding and genetic gain in selected populations. Ph.D. Diss., Wageningen Univ., Wageningen. Universal Press, Veenendaal, The Netherlands.

Bijma, P. and Woolliams, J.A., 2000. Prediction of rates of inbreeding in populations selected on best linear unbiased prediction of breeding value. Genetics, 156:361–373.

Bijma, P., Van Arendonk, J.A. and Woolliams, J.A. 2001. Predicting rates of inbreeding for livestock improvement schemes. Journal of Animal Science, 79:840-853.

Bright, G. 1991. Economic weights from profit equations: appraising their accuracy, Animal Production, 53:395-398.

Brisbane, J.R. and Gibson, J.P., 1994. Balancing selection response and rate of inbreeding by including genetic relationships in selection decisions. In Proceedings of the 5th World Congress on Genetics Applied to Livestock Production (Guelph) 19:135-138.

Brascamp, E.W., 1984. Selection indices with constraints. Anim. Brdg. Abs., 52:645-655.

Brisbane, J.R. and Gibson, J.P., 1995. Balancing selection response and rate of inbreeding by including genetic-relationships in selection decisions. Theoretical and Applied Genetics, 91:421–431.

Bulmer, M.G., 1971. The effect of selection on genetic variability. Am. Naturalist 105:201-211.

Clark, S.A., Kinghorn, B.P., Hickey, J.M. and van der Werf, J.H.J., 2013. The effect of genomic information on optimal contribution selection in livestock breeding programs. Genetics Selection Evolution 2013, 45:44

Cunningham, E.P., 1979. Definition, measurement and recording in the genetic improvement of livestock, Proceedings of the Australian Association of Animal Breeding and Genetics, 1:73-80.

Cunningham, E.P., Moen, R.A. and Gjedrem, T., 1970. Restriction of selection indexes, Biometrics, 26:67-74.

Daetwyler, H.D., Villanueva, B., Bijma, P. and Wooliams, J.A., 2007. Inbreeding in genome-wide selection. Journal of Animal Breeding and Genetics. 124: 369-376.

Ducrocq V. and Quaas R.L., 1988. Prediction of genetic response to truncation selection across generations. *Journal of Dairy Science* 71:2543-2553.

Falconer D.S. and Mackay T.F.C., 1996. *Introduction to Quantitative Genetics*. (Fourth edition), Longman.

Fernandez, J., Toro, M.A. and Caballero, A., 2001. Practical implementation of optimal management strategies in conservation programs: a mate selection method. *Animal Biodiversity and Conservation*, 24:1-7

Fowler, V.R., Bichard, M. and Pease, A., 1976. Objectives in pig breeding, *Animal Production*, 23:365-387.

Gibson J., 2005. Introduction to breeding objectives. Armidale Animal Breeding Summer Course 2005.

Gibson, J. and Dekkers, J., 2009. *Design and Economics of Animal Breeding Strategies*. University of New England - Animal Genetics.

<http://www.une.edu.au/tigb/Gibson-book/index.php>

Goddard, M.E., 1998. Consensus and debate in the definition of breeding objectives. *Journal of Dairy Science*, 81 (Suppl. 2): 6-18.

Goddard, M.E. and Howarth, J.M., 1994. In *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*. Guelph, 7 - 12 August 1994. 18:306-309.

Gondro, C. and Kinghorn, B.P., 2007. Solving Complex Problems With Evolutionary Computation. *Proc. Assoc. Advmt. Anim. Breed. Genet.* 17: 272-279.

Gondro, C. and Kinghorn, B.P., 2010. Application of Evolutionary Algorithms to solve complex problems in Quantitative Genetics and Bioinformatics. Armidale Animal Breeding Summer Course 2010. University of New England. Course notes.

Graser, H.U., Nitter, G. and Barwick, S.A., 1994. Evaluation of advanced industry breeding schemes for Australian beef cattle. II. Selection on combinations of growth, reproduction and carcass criteria. *Australian Journal of Agricultural Research*, 45:1657–69.

Graser, H.U., 2006. Optimising Breeding Program Design. Armidale Animal Breeding Summer Course 2006.

Grundy, B., Villanueva, B. and Wooliams, J., 1998. Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development. In *Proceedings of the 5th World Congress in Genetics Applied to Livestock Production*, 25:255–358.

Grundy, B., Villanueva, B. and Wooliams, J., 2000. Dynamic selection for maximising response with constrained inbreeding in schemes with overlapping generations. *Animal Science*, 70:373–382.

Grundy, B. and Hill, W.G., 1993. A method for reducing inbreeding with best linear unbiased prediction. *Animal Production*, 56:427.

Harris, D., 1998. Livestock improvement: Art, science, or industry. *Journal of Animal Science*, 76:2294–2302.

Harris, D.L., Stewart, T.S. and Arboleda, C.R., 1984. Animal breeding programs: a systematic approach to their design, AAT-NC-8. ARS, USDA, Peoria, IL.

Hayes, B.J., Shepherd, R.K. and Newman, S., 1998. A tactical approach to improving long-term response in across-breed mating plans. *Proceedings of the 6th World Congress on Genetics Applied to Livestock Production*, Armidale, vol. 23, pp. 439-442.

Henderson, C.R., 1973. Sire evaluation and genetic trends. *Proc. Anim. Breed. Genet. Symp. in Honor of Dr. J.L. Lush: 10*. American Society for Animal Science and American Dairy Science Assoc., Champaign, Illinois.

Henderson, C.R., 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics*, 32:69–83.

Henderson, C.R., 1984. *Applications of linear models in animal breeding*. University of Guelph Press, Guelph, Canada.

Henderson, C.R., 1990. Accounting for selection and mating biases in genetic evaluations. In Gianola, D. and Hammond, K., editors, *Statistical Methods for Genetic Improvement in Livestock*, pages 413–436. Springer-Verlag, Berlin.

Henderson, C.R., Kempthorne, O., Searle, S. and Vonkrosigk, C., 1959. The estimation of environmental and genetic trends from records subject to culling. *Biometrics*, 15:192–218.

Henryon, M., Sørensen, A.C. and Berg, P., 2009. Mating animals by minimising the covariance between ancestral contributions generates less inbreeding without compromising genetic gain in breeding schemes with truncation selection. *Journal of Animal Science*. 2009 Oct;3(10):1339-46

Hinrichs, D., Wetten, M. and Meuwissen, T.H.E., 2006. An algorithm to compute optimal genetic contributions in selection programs with large numbers of candidates. *Journal of Animal Science*, 84:3212-3218.

Hovenier, R., Brascamp, E.W., Kanis, E., van der Werf, J.H.J. and Wassenberg, A.P.A.M., 1993. Economic values of optimum traits: the example of meat quality in pigs, *Journal of Animal Science*, 71:1429-1433.

James, J.W., 1987. Determination of optimal decision policies. *Journal of Animal Breeding and Genetics*. 104, 23-27.

Kempthorne, O. and Nordskog, A.W. 1959. Restricted selection indices, *Biometrics*, 15:10-19.

Keller, D.S. and Gibson, J.P., 1990. Linear programming as a general method for solving unrestricted and restricted selection indexes. *Journal of Dairy Science* 73 (Suppl. 1):255 (Abstr.)

Kinghorn, B.P., 2000. The tactical approach to implementing breeding programs. Chapter 22 in “Animal Breeding – Use of New Technologies”,

Kinghorn, B.P., van der Werf, J.H.J. and Ryan, M. (eds.). The Post Graduate Foundation in Veterinarian Science of the University of Sydney. ISBN 0 646 38713 8. Pages 291-308.

Kinghorn, B.P., 2000. <http://metz.une.edu.au/~bkinghor/desire.htm>

Kinghorn, B.P., 2005. Desired gains and other desired outcomes. Armidale Animal Breeding Summer Course 2005.

Kinghorn, B.P., 2011. A grouping algorithm for efficient constrained mate selection. *Genetics Selection Evolution* 2011, 43:4. <http://www.gsejournal.org/content/43/1/4>

Kinghorn, B.P. and Kinghorn A.J., 2009. Pedigree Viewer Notes.

Kinghorn, B.P. and Kinghorn A.J., 2009. Pedigree Viewer system (<http://www-personal.une.edu.au/~bkinghor/pedigree.htm>)

Kinghorn, B.P. and Shepherd, R.K. 1994. A tactical approach to breeding for information-rich designs. In Proceedings of the 5th World Congress on Genetics Applied to Livestock Production. Guelph, 7 - 12 August 1994. 18: 255-261.

Kinghorn, B.P. and Shepherd, R.K., 1999. Mate selection for the tactical implementation of breeding programs. *Assoc. Advmt. Anim. Breed. Genet.* 13:130-133.

Kinghorn, B.P., Meszaros, S.A. and Vagg, R.D., 2001. Progeny Explorer. *Assoc. Advmt. Anim. Breed. Genet.*

Kinghorn, B.P., Meszaros, S.A. and Vagg, R.D., 2002. Dynamic tactical decision systems for animal breeding. *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*. 33: 179-186.

Koots, K.R. and Gibson, J.P., 1998. Economic values for beef production traits from a herd level bioeconomic model. *Canadian Journal of Animal Science*, 78:29-45.

Leitch, H.W., Smith, C., Burnside, E.B. and Quinton, M., 1994. Genetic Response and Inbreeding with Different Selection Methods and Mating Designs for Nucleus Breeding Programs in Dairy Cattle. *Journal of Dairy Science* 77, 1702-1718.

Lewis, R.M. and Simm, G., 2000. Selection strategies in sire referencing schemes in sheep. *Livestock Production Science*, 67:129–141.

Mayer, D.G., Kinghorn, B.P. and Acher, A.A., 2005. Differential evolution – an easy and efficient evolutionary algorithm for model optimisation. *Agricultural Systems* 85:315-328.

Meszaros, S.A., Banks, R.G., van der Werf, J.H.J. and Goddard, M., 1999. The use of genetic algorithms for optimising age structure in breeding populations when inbreeding depresses genetic gain through effects on reproduction, *Animal Science* 68:449-457.

Meuwissen, T.H.E., 1989. A deterministic model for the optimisation of dairy cattle breeding based on BLUP breeding value estimates. *Animal Production*, 49:193–202.

Meuwissen, T.H.E., 1997. Maximizing the response of selection with a predefined rate of inbreeding. *Journal of Animal Science*, 75: 934-940

Meuwissen, T.H.E., 2002. GENCONT: An operational tool for controlling inbreeding in selection and conservation schemes. In *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*. 28:769-770

Meuwissen, T.H.E. and Sonesson, A.K., 1998. Maximizing the response of selection with predefined rate of inbreeding: Overlapping generations. *Journal of Animal Science*, 76: 2575–2583.

Meuwissen, T.H.E. and Woolliams, J.A., 1994. Maximizing genetic response in breeding schemes of dairy-cattle with constraints on variance of response. *Journal of Dairy Science*, 77:1905–1916.

Meyer, K. and Smith, C., 1990. Comparison of theoretical and simulated equilibrium genetic response rates with progeny testing in dairy-cattle. *Animal Production*, 50:207–212.

Misztal, I., 1999. BLUPF90 - a mixed model program in Fortran 90. *Animal and Dairy Science*, University of Georgia.

Monti, A., 2004. Multiple trait evaluation in Angus beef cattle. Master thesis, University of New England.

Newman, S., Kremer, V.K. and Kinghorn, B.P., 2009. Mate selection strategies.—Balancing genetic gain and diversity in livestock improvement programs. *Proceedings of the National Breeders Roundtable*.

Nitter, G., Graser, H.U. and Barwick, S., 1994. Evaluation of advanced industry breeding schemes for Australian beef cattle. I. Method of evaluation and analysis for an example population structure. *Australian Journal of Agricultural Research*, 45:1641–1656.

Phocas, F., Bloch, C., Chapelle, P., Becherel, F. and Menissier, G.R.F., 1998. Developing a breeding objective for a French purebred beef cattle selection programme. *Livestock Production Science*, 57:49–65.

Pong-Wong, R. and Woolliams J.A., 2007. Optimisation of contribution of candidate parents to maximize genetic gain and restricting inbreeding using semi-definite programming. *Genetics Selection Evolution* 39:3-25

Ponzoni, R.W., 1982. Breeding objectives in sheep improvement programmes, *Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production*, 6:619-634.

Quinton, M. and Smith, C., 1995. Comparison of evaluation-selection systems for maximizing genetic response at the same level of inbreeding. *Journal of Animal Science*, 73:2208-2212.

Rendel, J.M. and Robertson, A., 1950. *J. Genet.* 50 : 1-8.

Segelke, D., Reinhardt, F., Liu, Z. and Thaller, G., 2014. Prediction of expected genetic variation within groups of offspring for innovative mating schemes. *Genetics Selection Evolution* 2014, 46:42.

Shultz, F.T., 1990. Formulation of breeding objectives for poultry meat production. Proceedings of the 3rd World Congress on Genetics Applied to Livestock Production, 10:215-227.

Sivarajasingam, S., 1995. A method to estimate economic weights for traits of disease resistance in sheep. Proceedings of the Australian Association of Animal Breeding and Genetics, 11:65-69.

Shepherd, R.K. and Kinghorn, B.P., 1998. A tactical approach to the design of crossbreeding programs. In Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, volume 25, pages 431–438. Armidale, Australia.

Schneeberger, M., Barwick, S.A., Crow, G.H. and Hammond, K., 1992. Economic indexes using breeding values predicted by BLUP. Journal of Animal Breeding and Genetics, 109:180–187.

Storn, S. and Price, K., 1997. Dr. Dobbs Journal 264: 18-24.

Storn, S. and Price, K., Differential Evolution Homepage (<http://www.icsi.berkeley.edu/~storn/code.html>)

Tallis, G.M. 1962. A selection index for the optimum genotype, Biometrics, 18:120-122.

Toro, M. and Pérez-Enciso, M., 1990. Optimization of selection response under restricted inbreeding. Genetics Selection Evolution 22, 93:107

Toro, M. and Sileo, L., 1992. An alternative to restricted BLUP based on mate selection. *Livestock Production Science*, 32:181–187.

van der Werf, J.H.J., 2005. Genetic change of multiple traits. *Armidale Animal Breeding Summer Course 2005*.

van der Werf, J.H.J., 2005. Optimising design of breeding programs. *Armidale Animal Breeding Summer Course 2005*.

van der Werf, J.H.J., 2007. Design of Breeding Programs. GENE522 notes. University of New England.

van der Werf J.H.J. et al., 2009. *Adaptation and Fitness in Animal Populations*, Springer Science+Business Media B.V.

Verrier, E., Colleau, J.J. and Foulley, J.L., 1993. Long-term effects of selection based on the animal model BLUP in a finite population. *Theoretical and Applied Genetics* 87:446-454.

Villanueva, B., Bijma, P. and Woolliams, J.A., 2000. Optimal mass selection policies for schemes with overlapping generations and restricted inbreeding. *Genetics Selection Evolution*, 32(4):339–355.

Weigel, K.A., 2001. Controlling inbreeding in modern breeding programs. *Journal of Dairy Science* 84(E. Suppl.):E177-E184.

Woolliams, J.A. and Bijma, P., 2000. Predicting rates of inbreeding in populations undergoing selection. *Genetics*, 154:1851–1864.

Woolliams, J.A., Bijma, P. and Villanueva, B., 1999. Expected genetic contributions and their impact on gene flow and genetic gain. *Genetics*, 153:1009–1020.

Woolliams, J.A. and Thompson, R., 1994. A theory of genetic contributions. In *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, 25:127-134. 1994.

Wray, N.R. and Hill, W.G., 1989. Asymptotic rates of response from index selection. *Animal Production*, 49:217–227.

Wray, N.R. and Goddard, M.E., 1994. Increasing long-term response to selection. *Genetics Selection Evolution* 26:431-459

Yamada, Y., Yokouchi, K. and Nishida, A., 1975. Selection index when genetic gains of individual traits are of primary concern, *Japanese Journal of Genetics*, 50:33-41.