Non-additive genetic effects in animal selection

Marek ŁUKASZEWICZ
Institute of Genetics and Animal Breeding, Polish Academy of Sciences, Jastrzębiec, Poland

Abstract. Genetic evaluation of purebred farm animals has been carried out for about half a century, employing additive approximation to describe the genetic background. An evaluated animal has been attributed a single breeding value for each trait of the breeding goal. The predicted additive genetic value of an animal equals the average breeding value of its parents. Although the selection based on the additive approach has proved successful, there still is a possibility of increasing the reliability of the breeding value estimation by accounting for non-additive genetic effects of dominance and epistasis, disregarded in the additive model. In the non-additive model, the expected quality of the progeny equals the average of the parents plus an effect resulting from the interaction between the parents. In this case, the evaluated animal may have as many breeding values as there are possible candidates to mate to, for each trait. The dominance and epistatic effects have already been accounted for in selecting animals or populations for some crossbreeding plans (combining ability, heterosis, and recombination loss). Also, using crossbreds for the sake of the breeding value estimation of purebred animals requires removing the non-additive effects from the crossbred performance and distributing the additive component between the purebreds. Combining ability is more and more discussed as a factor for matings within breed to produce terminal progeny.

Key words: additive effects, animal selection, crossbreds, dominance effects, epistasis, purebreds.
Introduction

An extensive description of the problem of non-additive genetic effects in breeding value estimation was given by MISZTAL et al. (1995). Besides some background theory that paper discussed the feasibility and consequences of accounting for the non-additive effects, including inbreeding, in genetic evaluations of purebreds. The results of selection using the additive model suggest that this model is a good approximation of the real gene actions. However, it becomes evident that this approach, though efficient enough for producing a breeding stock which can be called "future progeny", may not suffice to yield highly productive terminal "present progeny". The present paper will mainly focus on the theory underlying the genetic phenomena being established upon the uniting of gametes and responsible for the performance of farm animals at both purebreeding and crossbreeding levels. Some implications to the biological description (statistical modelling) of animals’ performance under the non-additive influence are also discussed.

Models for one, two or more loci

Let us assume that a trait is determined by only one autosomal locus. The genotype effect $g_{ij}$ of an animal with allele $i$ from gamete 1 and allele $j$ from gamete 2 can be divided as follows (e.g. BULMER 1985):

\[ g_{ij} = a_i + a_j + d_{ij} \]

where: $a_i$ and $a_j$ are the additive effects of alleles $i$ and $j$, and $d_{ij}$ is the dominance effect of alleles $i$ and $j$.

When a trait is affected by two loci with alleles $i$ and $j$ at locus A and alleles $k$ and $l$ at locus B, the genotype value of an individual with alleles $ijkl$ can be partitioned:
where: $aa_{ij}$, ..., $aa_{jl}$ are two-allele epistasis, $ad_{ijkl}$, ..., $ad_{ij}$ are three-allele, additive-dominance epistasis, and $dd_{ijkl}$ is four-allele, dominance-dominance epistasis.

The extension to a multi-locus model is straightforward. With an infinitesimal model, the overall genotype effect $(g)$ can be summed up as:

$$g = a + d + aa + ad + dd + aaa + aad + \ldots \ldots ,$$

where $a$ is the sum of additive effects over all the loci involved, $d$ is the sum of all the two-allele interaction effects within a locus, $aa$ is the sum of all the two-allele interaction effects between two loci, $ad$ is the sum of all the three-allele interaction effects involving two loci, $dd$ is the sum of all the four-allele interaction effects involving two loci, $aaa$ is the sum of all the three-allele interaction effects involving three loci, etc. In the absence of inbreeding, these effects are independent, and their variances can be written as:

$$\text{var}(a) = \sigma_a^2, \text{var}(d) = \sigma_d^2, \text{var}(aa) = \sigma_{aa}^2, \text{var}(ad) = \sigma_{ad}^2, \text{var}(dd) = \sigma_{dd}^2, \text{var}(aaa) = \sigma_{aaa}^2, \ldots$$

Since there are infinitely many effects and the total genotype variance is finite, most of the variances must be close to zero. It is generally presumed that variances of traits most selected for decrease with the increasing gene interaction order, and that the additive variance is the largest, justifying the popularity and effectiveness of the additive genetic model.

A knowledge of covariances between relatives is the must for the genetic evaluation of livestock. For animals $x$ and $y$ with coefficient of the additive relationship $\alpha$ and coefficient of the dominance relationship $\delta$, the covariance between genotype values of $x$ and $y$ is (e.g. CHANG 1988):

$$\text{cov}(x, y) = \alpha_{xy} \sigma_a^2 + \delta_{xy} \sigma_d^2 + \alpha_{xy} \sigma_{aa}^2 + \alpha_{xy} \delta_{xy} \sigma_{ad}^2 + \delta_{xy} \sigma_{dd}^2 + \alpha_{xy} \delta_{xy} \sigma_{aaa}^2 + \ldots$$

Such coefficients for popular types of relationships are given in Table 1.

<table>
<thead>
<tr>
<th>Relationship</th>
<th>additive relationship</th>
<th>dominance relationship</th>
<th>additive-additive relationship</th>
<th>additive-dominance relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clone</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Full-sib</td>
<td>0.5</td>
<td>0.25</td>
<td>0.25</td>
<td>0.125</td>
</tr>
<tr>
<td>Parent-offspring</td>
<td>0.5</td>
<td>0</td>
<td>0.25</td>
<td>0.0</td>
</tr>
<tr>
<td>3/4-sib*</td>
<td>0.3125</td>
<td>0.0625</td>
<td>0.0977</td>
<td>0.0195</td>
</tr>
<tr>
<td>Half-sib</td>
<td>0.25</td>
<td>0</td>
<td>0.0625</td>
<td>0</td>
</tr>
</tbody>
</table>

*The 3/4-sibs are animals who have one common parent and a grandparent of the other parent
If a given relationship between animals does not contain a certain gene action, then such a relationship is useless for the estimation of variance component or prediction of breeding values for that particular gene action. Further, the smaller is the coefficient of relationship, the less precise is the estimation or prediction. As MISZTAL et al. (1995) indicated, coefficients for higher order epistatic effects are decreasing, so those effects are also increasingly more difficult to estimate. Relationships amongst clones, full-sibs and 3/4-sibs are useful for estimating the dominance variance and/or predicting the dominance effects, but parent-progeny and half-sib relationships are not (Table 1).

**Inbreeding**

When a population is subjected to inbreeding, variances of different effects change, covariances among those effects are created, and inbreeding depression decreases the performance of animals (KENNEDY et al. 1988). Changes in covariances between the additive and dominance effects can be ignored with only little loss in accuracy, provided that the model includes a covariable for the inbreeding coefficient (DE BOER, HÖSCHELE 1993):

$$g = f_{s,d} \Delta + a + d + \cdots,$$

where: \(f_{s,d}\) is the coefficient of inbreeding for progeny of sire \(s\) and dam \(d\), and \(\Delta\) is the estimate of inbreeding depression.

**Implications of ignoring the non-additive effects**

If the non-additive effects in the model are disregarded, the other effects estimated with such a model are biased to a varying degree. The following repeatability model used for data with repeated observations per animal illustrates the problem (MISZTAL et al. 1995):

$$y = X \beta + Z (a + p + f\Delta + d + aa + ad) + e$$

The animal-specific permanent environment effect \(p\) is a non-genetic part of the animal's performance variance. Out of the omitted non-additive effects, the dominance effect \(d\) becomes mostly a part of the permanent environment effect, and the epistatic effects \(aa\) and \(ad\) are confounded with both the additive
and the permanent environment effects. $X$ and $Z$ are design matrices and $\beta$ is a vector of fixed effects.

In a single record model, where the permanent environment effect is absent, effects that would otherwise be confounded with the permanent environment effect, become a part of the residual effect (MISZTAL et al. 1995):

$$ y = X\beta + Z(a + f\Delta + d + aa + ad) + e $$

Consequences of ignoring non-additive effects are different for various groups of animals. They are the largest for animals whose relationships contain large non-additive covariances (VANRADEN, HÖSCHELE 1991), for example clones and full-sibs. Even though the number of animals of such relationships may be small compared to the population as a whole, they may be the elite of the population, and their influence on the population could be disproportionately greater. For example, cloning in dairy cattle is restricted only to the best animals. Similarly, numerous full-sibs are results of embryo transfers – a procedure reserved for future sires and bulldams (MISZTAL et al. 1995).

**Approximate variance components of the non-additive effects**

The prediction of a breeding value assumes knowledge of the variance components for particular gene actions we are interested in. We can check if it is worthwhile to account for the non-additive effects when predicting animals' breeding value, by using different instances of the additive model. The upper limit of the total non-additive variance for animals with repeated records can be calculated from the repeatability model. Let us assume that non-additive effects ignored in the model become a part of the permanent environment effect. Repeatability ($r^2$) is defined as the ratio of the sum of the genetic and permanent environment variances to the phenotypic variance (e.g. MISZTAL et al. 1995):

$$ r^2 = \left( \frac{\sigma_a^2 + \sigma_{pe}^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2} \right), $$

where $\sigma_{pe}^2$ is the variance of permanent environment effect. Then, that variance consists of the non-genetic ($\sigma_n^2$) and non-additive ($\approx \sigma_a^2 + \sigma_{aa}^2 + \sigma_{ad}^2 + \sigma_{dd}^2 + \ldots$) variances:

$$ \sigma_{pe}^2 = \sigma_n^2 + \sigma_a^2 + \sigma_{aa}^2 + \sigma_{ad}^2 + \sigma_{dd}^2 + \sigma_{aaa}^2 + \ldots $$

Assuming that heritability estimated with the animal model contains only the additive variance in the numerator:

$$ h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2}, $$
the upper limit of the non-additive variance relative to the phenotypic variance is:

\[
\frac{\sigma^2_d + \sigma^2_{aa} + \sigma^2_{ad} + \sigma^2_{dd} + \sigma^2_{aaa} + \ldots}{\sigma^2_a + \sigma^2_{pe} + \sigma^2_e} \leq r^2 - h^2.
\]

In the case of milk production traits, repeatability is about 60% higher than heritability. Thus, all the non-additive variances combined account for at most 60% of the additive variance. The estimates based on the repeatability model are only approximate because some non-additive variances in the additive model are estimated as additive variance (UIMARI, KENNEDY 1990), especially when populations contain many full-sibs, e.g., in swine or poultry. In fact, the overall non-genetic variance may be even smaller since the term \( \sigma^2_{pe} \) includes \( \sigma^2_n \), too.

**Genetic merit of progeny**

In the additive model the genetic merit of a future progeny of sire \( s \) and dam \( d \) can be written as:

\[
g = \frac{1}{2} a_s + \frac{1}{2} a_d.
\]

Each parent is assumed to have one genetic evaluation (breeding value) for each trait and the number of breeding values would easily fit in a catalog.

In the non-additive model, the genetic quality of a progeny can be presented as (MISZTAL et al. 1995):

\[
g = a_s^* + a_d^* + n_{s,d}
\]

effects specific to each parent (averaged over possible mates)

or

\[
g = a_d^* + a_{s,d}
\]

effect of sire specific to each dam.

According to the last formula, one sire (or parent in general) can have multiple breeding values, one for each mate. In large populations the usual number of evaluated animals is big enough to make the handling and publishing of the proofs impossible.

A solution to that problem is selecting animals only on the basis of the additive effect, hoping that the presence of the non-additive effects in the model increases the accuracy of the additive predictions. Unfortunately, as UIMARI and KENNEDY (1990) have shown, such an attitude overestimates the genetic value of future progeny and results in a lesser genetic gain than in the classical additive model without the non-additive effects. On the other hand, correction for inbreeding in the model may favor highly inbred individuals, and may inflate prediction of the additive effect as well (LAWLOR et al. 1993). Yet, there is some evidence that accounting for inbreeding in the model does not influence significantly the response to selection (MEHRABANI-YEGANCH et al. 2000). MISZTAL et al. (1995)
conclude that non-additive effects should not be accounted for in the model unless they are used in selection.

**Breeding value vs. selection horizon**

If we define breeding value as a quality measure of animals as parents of the following generation, we face a dual situation reflecting the use of additive or non-additive approaches. Parents transmit their genetic value upon progeny through the gametes. It means that they pass a sample of their alleles and combinations of the alleles within a gamete. In other words, they pass a sample of their additive quality and additive epistatic interaction of alleles present in the gamete. Upon uniting of gametes a new quality emerges – interactions of alleles within locus, and between loci and gametes. Parents who usually produce excellent gametes, do not necessarily produce the best progeny, due to their possibly inferior combining ability.

To maximize the total merit in the immediate generation, one would select for the total genetic merit:

$$\max (fA + at + d + aa + ...)$$

where the non-additive effects present in progeny are maximized through selection of parents of a high combining ability. In contrast to numerous breeding values attached to each possible mate, which can be called “specific combining ability”, the “general combining ability” is averaged over all the actual matings or combinations of genes of a candidate, traced with the aid of adequate relationship (covariance) matrices.

To maximize the total genetic merit in future generations, one would select only for the additive effect, perhaps with an adjustment for inbreeding to minimize inbreeding depression:

$$\max (a + fA)$$

The above would require parents producing best quality gametes.

In certain species, e.g. in dairy cattle, one sex is used predominantly for production and the other one for reproduction. HENDERSON (1989) suggested that if the non-additive effects were used in selection, it might be beneficial to maximize the current generation merit for the “production” sex and the future generation merit for the “reproduction” sex.

While the epistatic effects are still more a theoretical problem, due to a low number of adequate relationships and computational problems, evaluation of the dominance effects is on the best way to become an everyday routine. Along with finding the method for fast inversion of the dominance relationship matrix (HÖSCHELE, VANRADEN 1991), and revealing significant dominance effects on farm animals’ performance (e.g. TEMPELMAN, BURNSIDE 1990, KLAASKATE et al. 1994, RODRIGUEZ-ALMEIDA et al. 1995, CULBERTSON et al. 1997,
MISZTAL et al. 1997, GENGLER et al. 1998), mate selection strategies to exploit the dominance variance and combining ability of potential parents have been searched for (e.g. DE STEFANO, HÖSCHELE 1992, VARONA, MISZTAL 1999, HAYES, MILLER 2000).

**Non-additive effects in crossbreeding**

In fact, the non-additive effects were first, and intentionally, exploited in selecting populations for various crossbreeding schemes. The notions of “heterosis” and “recombination loss” associated with crossbreeding, reflect dominance and epistatic gene actions at the interpopulational level.

**Heterosis vs. dominance**

Dominance is the effect of interaction between alleles within one locus, regardless the direction of the effect. However, there is some evidence that the evolution tends to create and maintain beneficial dominance combinations, eliminating undesired alleles (FISHER 1931). That phenomenon is called “directional dominance” and is the main cause of inbreeding depression and heterosis.

Let us consider a quantitative trait locus with two alleles present in a population with frequencies of $p$ and $q$, while $\mu$ is the average effect of the other loci determining that trait. The values of three possible genotypes can be presented (FALCONER 1981):

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Value</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1A_1$</td>
<td>$&lt; \mu + a$</td>
<td>$p^2$</td>
</tr>
<tr>
<td>$A_1A_2$</td>
<td>$&lt; \mu + d$</td>
<td>$2pq = H$</td>
</tr>
<tr>
<td>$A_2A_2$</td>
<td>$&lt; \mu - a$</td>
<td>$q^2$</td>
</tr>
</tbody>
</table>

where $a$ is the deviation of homozygotes and $d$ is the deviation of heterozygotes from the mean.

The population mean amounts to:

$$M = p^2(\mu + a) + 2pq(\mu + d) + q^2(\mu - a) = \mu + (p - q)a + Hd.$$  

The first two terms in the formula do not depend on heterozygosity ($H$). The third term is the product of heterozygosity and the dominance effect related to that locus. Assuming that this effect differs from zero, the change in the genetic mean as-
Non-additive genetic effects in animal selection

Non-additive genetic effects in animal selection associated with inbreeding or heterosis is a linear function of heterozygosity. Extension to a higher number of loci requires directional dominance and lack of epistatic interaction assumption.

The expected performance of crossbreds may be presented as (CUNNINGHAM, SYRSTAD 1987):

\[ M = \mu + QA + Hd. \]

In that formula \( \mu \) stands for the mean of one of the parental populations, \( Q \) is the proportion of the other population in crossbred animals, \( a \) is the additive difference between populations, and \( d \) is the heterosis effect found in the \( F_1 \) generation. This approach is based on the definition of heterozygosity from the breed-of-allele-point-of-view (KINGHORN 1982). In such a case, the \( F_1 \) generation is fully heterozygotic as both alleles at a locus originate from different breeds.

To predict the performance of different crosses between two breeds we should know the values of \( a, d, Q, \) and \( H. H \) and \( Q \) are obtained from the structure of matings, while \( a \) and \( d \) have to be estimated on the actual data.

If we assume two populations \( P_1 \) and \( P_2 \), with the additive genetic difference of \( a \) and heterosis \( d \) equal to \( F_1 - 0.5(P_1 + P_2) \), then the expected performances of various groups are:

\[
\begin{align*}
P_1 &= \mu, \\
P_2 &= \mu + a, \\
F_1 &= \mu + 0.50a + d, \\
F_2 &= \mu + 0.50a + 0.5d, \\
B_1 &= \mu + 0.25a + 0.5d, \\
B_2 &= \mu + 0.75a + 0.5d.
\end{align*}
\]

The additive-dominance model proved to be useful in many instances in different species. Yet, SHERIDAN (1980, 1981) showed that this model did not suffice to explain certain observations in laying hens and dairy cattle.

**Heterosis vs. epistasis**

Genetic structure established during purebred selection, with favorable combinations of non-allelic genes within a gamete, may undergo recombination when producing advanced generations of crossbreds. As a result, crossbreeding effects caused by recombination may appear unfavorable, decreasing the performance of crossbreds. DICKERSON (1969) called this phenomenon "recombination loss".

It can be concluded from the covariance between parent and offspring (\( \text{cov}_{po} = 0.5\sigma_a^2 + 2.25\sigma_{aa}^2 + 1.25\sigma_{aaa}^2 + \ldots \)) that selection affects only the additive genetic interactions between the non-allelic genes. Further, the interactions between a greater number of loci will become negligible as the coefficients of relationship

The expected merit of a group of crossbreds can be described with the following general model:

\[ E = \sum_{i=1}^{n} w_i a_i + [c_{w}d + c_{bg}e + c_{wg}e] \]

where: \( E \) = expected performance of crossbreds, \( n \) = number of breeds, \( w_i \) = proportion of \( i \)-th breed in the crossbred genotype, \( a_i \) = additive effect of \( i \)-th breed, \( c_{w} \) = probability that two alleles at one locus originate from different breeds, \( c_{bg} \) = probability that two alleles at different loci and gametes originate from different breeds, \( c_{wg} \) = probability that two alleles within a gamete are of different breed origin, \( d \) = dominance effect, \( e \) = epistasis effect.

The expression in brackets is simplified and describes the general combining ability rather than the specific ones.

Kinghorn combined the coefficients \( c_{bg} \) and \( c_{wg} \)

\[ E = \sum_{i=1}^{n} w_i a_i + [c_{w}d + (c_{bg} + c_{wg}) / 2e] \]

In the Dickerson's model accepted is fact that always \( c_{w} = c_{bg} = c \):

\[ E = \sum_{i=1}^{n} w_i a_i + [(d + e) + c_{wg}e] \]

The heterosis effect is composed of the dominance and the epistasis effects, and \( e \) is the recombination loss. With this model heterosis is estimated without bias caused by other genetic effects in the classical comparison of \( F_1 \) generation performance with the average performance of parental populations.

Having estimated the effects of heterosis and recombination loss for various sets of breeds, one can decide which breeds would best suit the needs. To improve the effectiveness of a crossbreeding plan even more, the selection of purebred animals for the best combining ability can be performed.

**Using crossbred performance for breeding value estimation of purebreds**

With small populations, when many genes of pure breeds are present in crossbred animals (swine, poultry), we can be interested in using the information contained in crossbreds to increase the accuracy of purebred proofs. To be able to do it we must get rid of the non-additive effects (heterosis, recombination loss) from the crossbred performance and properly distribute the additive effects between the pure breeds. It appears that the efficiency of selection depends on the line contribution to the additive genetic variance of crossbreds. The loss of efficiency may amount to 13% when one line contributes only 11% of the variance and drops down to 0.2% with a 40% contribution (MIELENZ et al. 2000).
Conclusions

The non-additive effects were first exploited in selecting populations for various crossbreeding schemes. Along with the computer hardware and software development it has become possible to account for the non-additive effects in breeding value estimation of purebred animals to produce “future progeny” or to maximize production in “current progeny” by selecting parents for their combining ability. There are, however, some limitations to estimation of the non-additive effects connected mostly with the lack of adequate relationships in the evaluated populations.

REFERENCES


