Estimation of genotype × environment interaction for yield, health and fertility in dairy cattle

Mario P.L. Calus

Estimation of genotype × environment interaction for yield, health and fertility in dairy cattle

Promotor:

    Prof. Dr. ir. J.A.M. van Arendonk

    Hoogleraar Fokkerij en Genetica, Wageningen Universiteit

Co-promotor:

    Dr. ir. R.F. Veerkamp

    Clusterleider, Animal Sciences Group – Wageningen UR

Promotiecommissie:

Prof. Dr. ir. E.W. Brascamp	(Wageningen Universiteit)
Dr. F.A. van Eeuwijk	(Wageningen Universiteit)
Prof. Dr. R.F. Hoekstra	(Wageningen Universiteit)
Dr. E.A. Mäntysaari	(Agricultural Research Centre – Finland)

M.P.L. Calus

Estimation of genotype × environment interaction for yield, health and fertility in dairy cattle

Estimation of genotype × environment interaction for yield, health and fertility in dairy cattle

Abstract

In dairy cattle breeding, health and fertility traits have recently been included in a large number of national breeding goals. The effectiveness of breeding decisions and management changes to improve health and fertility possibly interact through genotype × environment interaction (G×E). G×E is a phenomenon in which different genotypes respond differently to changes in an environment. It can consist of the following effects: heterogeneous genetic variances across environments, genetic correlation of a trait expressed in different environments being less than 1.0 (reranking), and heterogeneous genetic correlations between traits across environments. In this thesis, G×E for health and fertiltity, as well as for yield, has been investigated using reaction norm models. In the reaction norm models, breeding values and genetic parameters were modeled as a function of an environmental descriptor using random regression. The dimensions of the model were expanded from linear random regressions to higher order random regressions, to include two parameters to define the environment, and to multivariate reaction norm models. Many environmental descriptors were investigated in this thesis, such as production level, farm size, average somatic cell score and calving interval, however, it appears that the herd parameters linked to nutrition and energy balance are most important for G×E. Significant G×E was detected in 86% of the situations for yield traits, but only in 14% of the situations for health and fertility traits, although significant reranking was found for SCS, mastitis and survival. Estimated G×E effects mainly consisted of heterogeneous genetic variances with limited reranking. As a result of heterogeneous variances in different traits, the relative importance of fertility compared to yield doubled across environments. Estimated G×E effects for SCS indicated more reranking of animals based on analysis of test day records, than on lactation averages. It was shown that selection for increased yield is expected to lead to increased environmental sensitivity for yield, while selection for better fertility is expected to lead to decreased environmental sensitivity for fertility. The models presented in this thesis can be used to account for the effect of herd environment on a trait and the relations between traits, and therefore enable to make accurate predictions of breeding values across environments.

Contents

Chapter 1

General introduction

M. P. L. Calus[1,2]

[1]Animal Sciences Group

P.O. Box 65, 8200 AB Lelystad, The Netherlands

[2]Wageningen Institute of Animal Sciences

Wageningen University

P.O. Box 338, 6700 AH Wageningen, The Netherlands

Chapter 1

GENERAL INTRODUCTION

The interest in breeding for health and fertility in dairy cattle has increased in the last decade. This is partly influenced by the trend towards more sustainable farming systems, but also due to the rapid increase in milk yield per cow. One-sided selection for increased milk yield predict higher risk of behavioural, physiological, and immunological problems for those animals, due to negative genetic correlations between yield traits and fertility, health, and metabolic traits (1988; Pryce et al., 1998; Rauw et al., 1998). This can result in an increase of on-farm costs (Jones et al., 1994), increased risks related to food safety (e.g., due to medicine use), reduced animal welfare, and societal concerns with regard to animal welfare. Two applied options to counter these side-effects of selection for increased milk yield, are: 1) improvement of management to account for the poorer genetic merit for health and fertility, and 2) inclusion of health and fertility in genetic selection with sufficient weight to ensure there is no reduction of the genetic level for health and fertility (Philipsson et al., 1994). Both options assume that management (the environment - E) and genetics (genotype - G), together add up to the phenotypic performance (P) of an animal: P = G + E. When management changes and genetic improvement interact, known as genotype × environment interaction (G×E), then phenotypic performance becomes: P = G + E + G×E. These G×E effects might be such, that effects of management changes and genetic improvement enlarge each other (G×E > 0), or, conversely, reduce each other (G×E < 0), and this is one reason why it is important to estimate the effects of G×E for health and fertility traits.

Making breeding decisions today, implies that one has to think which traits are economically important in the future (i.e., what is the breeding goal?). In the presence of G×E this becomes even more complex, because it is also important to consider the future herd environment and its relations with the breeding goal traits. Similarly, current breeding stock is used in many different environments, and data is collected also in many different herds. Therefore, in order to account for future and current herd management circumstances in animal breeding, the need exists to estimate G×E effects for health and fertility traits across a range of environments. However, most of the research into G×E has been geared towards yield traits and little is known about the importance of G×E effects for health and fertility traits.

Estimation of G×E effects associated with continuous levels of herd environment, was described in animal breeding, by the use of covariance functions, either in a two-step procedure (Veerkamp and Goddard, 1998) or in random regression models (Calus et al., 2002; Kolmodin et al., 2002). These models are also known as reaction norm models and find their origin in evolutionary biology (Woltereck, 1909). The simplest reaction norm model estimates breeding values of an animal as a linear function of their environmental values, across a range of environments. Reaction norm models can further be extended, to estimate G×E, 1) for a single trait as non-linear function of environmental values, or as a function of

more than one environmental values, and 2) for multiple traits, where not only the variances of the traits, but also the covariances between traits, depend on the environmental value. The application of random regression models enables a continuous definition of environments, rather than arbitrary chosen groups, as most often is applied when estimating G×E. Therefore it is expected that these models allow better estimation of G×E effects. However at the start of this study only a few studies were available estimating G×E with a random regression model, and therefore developing these models for G×E estimation is also of importance for this thesis.

AIM OF THIS THESIS

The aim of this thesis is to investigate the magnitude of genotype × environment interaction (G×E) for yield, health and fertility traits in dairy cattle using random regression models to estimate G×E.

OUTLINE OF THIS THESIS

In chapter 2, a review of the literature with regard to definition, estimation and importance of G×E is presented with main emphasis on G×E in dairy cattle. In chapter 3, the application of reaction norm models to estimate G×E for milk yield traits in relation to fourteen different environmental parameters is described. In chapter 4, the effects of possible confounding between environmental parameters and the analyzed trait were investigated. In chapter 5, the association of several descriptors of herd environment with phenotypic levels of fertility and health and G×E for fertility and health traits were investigated. In chapter 6, the reaction norm model was applied to milk fat yield and fat percentage on a test-day level, as a way to relate G×E for those traits to a metabolic disorder called milk fat depression. In chapter 7, G×E was estimated for somatic cell score based on test-day records, with a reaction norm model in three dimensions: bulk milk somatic cell count, stage of lactation, and the interaction of bulk milk somatic cell count and stage of lactation. In chapter 8, genetic correlations between milk production and health and fertility traits were estimated depending on herd environment, using a multi trait reaction norm model. Chapter 9, the general discussion, 1) discusses the main estimated G×E effects and their implications for breeding and management decisions in general, and on robustness of dairy cows and risks of high milk yield specifically, and 2) discusses the random regression methodology to estimate G×E, and compares it with structured antedependence models.

Chapter 2

Literature review

M. P. L. Calus

R. F. Veerkamp

Animal Sciences Group

P.O. Box 65, 8200 AB Lelystad, The Netherlands

Chapter 2

INTRODUCTION

Phenotypic performance of animals varies widely. Underlying mechanisms that affect phenotypic performance, such as physiology and adaptation of the animals, are due to both environmental and genetic factors. In animal husbandry, phenotypic performance is increased by improving management and selecting animals with desired genetic ability. An important question is to what extent environmental and genetic factors act additively or in an interactive manner and whether these genotype × environment interactions (G×E) lead to possible risks (i.e. decrease of phenotypic performance) or opportunities (i.e. increase of phenotypic performance). The aim of this study is to review 1) theory behind G×E, 2) estimation of G×E, 3) estimated levels of G×E, and 4) implications for the dairy industry.

GENOTYPE AND ENVIRONMENT

Definition of genotype

Genotype × environment interaction can be identified if phenotypic performance of at least two genotypes in at least two environments is considered (Mathur, 2002), which implies that both genotypes and environments need to be defined. Genotype can be defined as a unit, but also as a value of the genotype: a genotypic value. Genotypic units can be breeds (Kellaway and Colditz, 1975; Dillon et al., 2003a; Dillon et al., 2003b; Steinheim et al., 2004), crossbreds (Albers et al., 2002), groups of animals selected for a certain performance or genetic merit (Simm et al., 1994; Veerkamp et al., 1995), sires of animals (Tong et al., 1977; Meyer, 1987), animals (Berry et al., 2002), but also QTLs (Stratton, 1998), marker genotypes or genes.

Definition of environment

Genotype × environment interaction can only be identified, if at least two different environments are considered. Environment can be defined as a unit, but also as a continuous value of the environment. In animal husbandry, the most common environmental unit is herd. Each herd environment can be considered as a unique environment, affected by all possible characteristics and management decisions in that herd environment. However, using this definition, no inferences are made about the characteristics of the herd which makes understanding and comparison of different environments difficult. Sometimes distinct environments do have a specific meaning, without defining details, such as commercial versus nucleus herds (Merks et al., 1985). Defining environments as groups of herds with for instance a certain production level, not only reduces the number of distinct environments, but also makes comparison and ordering of environments easier. More precisely, each herd environment can be defined based on for instance average production level. All possible specific herd characteristics and combinations of herd characteristics can be used for this purpose. The benefit of using a single characteristic is that inferences can be made about the

relation between the considered herd characteristic and management, such as bulk milk somatic cell count related to clean versus dirty milking management practice (Barkema et al., 1999b). In literature, environments are defined as differences in climates (Mathur and Horst, 1994; Cienfuegos-Rivas et al., 1999; Costa et al., 2000; Kolmodin et al., 2004), different regions (Carabaño et al., 1990; Dodenhoff and Swalve, 1998; Kolmodin et al., 2004), specific differences between herds, such as low versus high concentrate levels (Veerkamp et al., 1994; Cromie, 1999; Pryce et al., 1999; Keady et al., 2001), organic versus conventional farms (Nauta et al., in press), housing system (Wicks and Leaver, in press; Buenger et al., 2001; Fatehi et al., 2003), management system (Boettcher et al., 2003; Fatehi et al., 2003), milking system (Mulder et al., 2004), average production level (Hill et al., 1983; De Veer and Van Vleck, 1987; Dong and Mao, 1990) or herd size (Kolmodin et al., 2002; Hayes et al., 2003). Based on a number of single characteristics, some sort of aggregate herd characterization can be defined, for instance by application of a canonical correspondence analysis (Faye et al., 1997), an applied factor analysis (Enevoldsen et al., 1996), or principal component analysis (Windig et al., 2005c). The principal components presented by Windig et al. (2005c), reflected for instance within-herd intensity of production, fertility and geographical positioning, and scale of dairy farms.

Defining environments on a geographical basis seems a logical thing to do, as weather conditions, soil and possibly social-economic perspectives change from region to region. Evidence has been found for heterogeneous genetic variances (i.e. scaling effects) for protein yield across different regions in Germany (Dodenhoff and Swalve, 1998) and genetic correlations for milk production traits between regions in the United States ranged from 0.93 to 0.99 (Carabaño et al., 1990). In small countries such as The Netherlands, estimated G×E across regions was however small (Van der Werf and Ten Napel, 1991). Due to the international character of current animal breeding and the interest in international breeding value estimation, the definition of different countries as different environments (Schaeffer, 1994) becomes increasingly more important, although it is questionable whether country boundaries are satisfactory to define environments (Weigel and Rekaya, 2000). A proposed solution, which might be more efficient and realistic, involves clustering of herds based on herd characteristics (Weigel and Rekaya, 2000; Zwald et al., 2001; Zwald et al., 2003).

Genotype × environment interaction

Usually, the phenotypic performance (**P**) of an animal is considered to be the sum of the value of its genotypic effect (**G**) and the value of an environmental effect (**E**): **P** = **G** + **E**. Consequently, the phenotypic variance (V_P) of a trait is equal to the sum of the genetic (V_G) and environmental variance (V_E): $V_P = V_G + V_E$, assuming that no covariance exists between G and E (Falconer and Mackay, 1996). In some situations the value of the genetic effect is affected by the environmental effect. This is a result of the ability of an animal to respond to

changes in the environment, also known as phenotypic plasticity or environmental sensitivity. More specifically, the phenomenon that environmental changes have different effects on different genotypes is known as G×E or differences in environmental sensitivity of genotypes (Falconer and Mackay, 1996).

Genotype × environment interaction results in three possible effects: 1) heterogeneity of genetic variances across environments (also known as scaling effects), 2) reranking of animals across environments based on estimated breeding values, and 3) heterogeneity of correlations between two or more traits across environments. The G×E is usually considered to be unimportant if only scaling effects and no reranking occurs. However, Namkoong showed that scaling effects of single traits can cause reranking of animals based on a composite index. Typically, estimated genetic correlations across environments are used to estimate the degree of reranking. In a situation where the genetic covariance between environments is not known, sometimes the correlation or rank correlation between estimated breeding values in both environments is used.

Genetic correlations between traits mainly result from pleiotropy, the characteristic of genes that they affect two or more traits (Falconer and Mackay, 1996). If the expression of some of the genes that affect traits, changes across environments, the genetic correlation between traits measured across environments may also change. This implies that the genetic correlation between traits can change due to changes in both the environment and the genotype (i.e. selection). From a biological perspective, the same conclusion can be derived, considering a situation where traits compete for limited available resources (Beilharz et al., 1993; Sölkner and James, 1994; Van der Waaij, 2004). Consider for example the trade-off between fertility and milk production in early lactation of dairy cows resulting from a negative energy balance due to lower energy intake than energy output (Berglund and Danell, 1987). The limitation of resources can be influenced by the genotype (i.e. the feed intake capacity of the cow is limiting) or the environment (i.e. the amount of feed available or nutritional value of the diet is limiting). Relaxing the limitation of resources will lead to a relaxation of the trade-off between traits (Van der Waaij, 2004) and therefore the genetic correlation between traits can change with changes in the genotype or changes in the environment.

METHODOLOGY TO ESTIMATE G×E EFFECTS
Animal experiments versus field data

In order to be able to estimate G×E, information of both genotypes and environments is required. On one hand one wants as detailed information as possible, but on the other hand a large number of observations is usually needed to be able to estimate G×E. Satisfying both conditions in one setting leads to enormous costs. If the argument of detailed information is most important, usually an animal experiment with typically a few dozen up to a few hundred

animals is applied. If a larger amount of data is most important, for instance for the purpose of variance component estimation, usually data available from field studies or milk recording schemes is used.

Comparison of mean performance across environments

Animal experiments in dairy cattle are expensive, but can be very illustrative. Typically, experiments are applied if required information is not available from routine milk recording data. Animals are usually split in a number of groups, based on level of milk production or genetic merit (Buckley et al., 2000; O'Connell et al., 2000). Some experiments run for a number of years and use two genetic lines: one has been selected for high production or breeding values and the other is a control line that has been selected for average production (Veerkamp et al., 1994; Veerkamp et al., 1995; Pryce et al., 1999). Environments are usually defined based on differences in feeding level or system (Veerkamp et al., 1994; Buckley et al., 2000; O'Connell et al., 2000; Keady et al., 2001). To refine the analysis, sometimes the discrete criteria to define genotypes is replaced by a continuous parameter, such as an estimated breeding value or a pedigree index (Veerkamp et al., 1995; Pryce et al., 1999). The phenotypic performance of the animals is within environment regressed on this parameter. Differences in regression coefficients across environments are then used as an indication for G×E.

Estimation of variance components for G×E in large data sets

The G×E can cause reranking of animals across environments, heterogeneous genetic variances across environments, (Lynch and Walsh, 1998) or heterogeneous genetic correlations between traits expressed in different environments. Traditionally, animal breeders are more concerned about reranking of animals than about differences in scale between environments because differences in scale do not affect the ranking of animals across environments for the considered trait. However, the magnitude of genetic variance affects genetic gain, and, therefore, the change in genetic variance across environments is important as well.

The most straightforward and computationally least demanding approach to estimate variance components for G×E is using a random sire×herd effect: $P = G + G×E + E$. This term measures the specific sire variance across herds (Robertson, 1959), and generally accounts for between 0 and 5% of the total phenotypic variance for yield traits (Dickerson, 1962; Tong et al., 1977; Meyer, 1987; Van der Werf and Ten Napel, 1991). The sire×herd method has the disadvantage that a large number of effects has to be estimated, while the number of informative animals per effect is limited and the additive genetic relation between sires is not considered in the estimation of G×E. Besides, the estimated breeding values do not include information of differences in performance of genotypes in different environments. Another

possibility to estimate G×E is to apply a multitrait model, i.e. one trait is considered to be a different trait in another environment: P = (G + G×E) + E. This method includes the interaction variance in each defined environment in the genetic variance and is often used to estimate G×E (Brotherstone and Hill, 1986; De Veer and Van Vleck, 1987; Boldman and Freeman, 1990; Carabaño et al., 1990), also for the purpose of international breeding value estimation (Schaeffer, 1994). A disadvantage of this method is that arbitrary grouping of environments is required and that the number of environments has to be kept small to limit the number of (co)variance components that has to be estimated. The most recently introduced method in animal breeding to estimate G×E implies defining a trait as a function of the environment: P = G(E) + E. The resulting model, usually referred to as reaction norm model, finds its origin in evolutionary biology (Woltereck, 1909) and describes the phenotypic performance of an animal as function of a continuous measurement of the environment. The environmental value is usually calculated as the mean phenotypic performance in a specific environment. The genetic effect is modeled by a covariance function (Kirkpatrick and Heckman, 1989) that models the environmental sensitivity of each genotype. The major benefit is that arbitrary grouping of environments is avoided, while the number of estimated variance components is limited. Applications in dairy cattle are recently described in a number of studies (Veerkamp and Goddard, 1998; Strandberg et al., 2000; Calus et al., 2002; Kolmodin et al., 2002).

Dealing with G×E in breeding value estimation

If breeding values are estimated in the presence of G×E, but without taking G×E into account, estimated breeding values will be biased. If scaling effects are not accounted for, animals in environments with high variance will have a higher chance to be selected. Correcting for G×E has been done in several ways. A sire × herd effect can be included in the model to account for G×E. Also, several methods have been proposed to account for heterogeneity of variance, either as pre-adjustment in the data (Boldman and Freeman, 1990; Wiggans and VanRaden, 1991; Weigel and Lawlor, 1994; Dodenhoff and Swalve, 1998) or as correction in the breeding value estimation model (Meuwissen et al., 1996). Including of G×E in breeding values has the advantage that account of G×E can be taken in selection decisions, whereas correction for heterogeneity assumes selection decisions for the average herd are applicable in all herd environments. Including G×E in the breeding value estimation can be done with a multitrait model or a reaction norm model. The reaction norm model, however, combines the advantages that no arbitrary grouping of environments is required while the number of estimated parameters is limited. On the other hand, in the reaction norm model some sort of inference about the continuous measure of the environment has to be made.

MAGNITUDE OF ESTIMATED G×E

Yield traits

For dairy cattle, hardly any significant G×E is reported in animal experiments where genotypes were defined based on genetic merit or selection on production, while environment was defined based on different diets (Veerkamp et al., 1994; Veerkamp et al., 1995; Buckley et al., 2000; Keady et al., 2001). However, regression of performance on pedigree index did show an interaction in some studies (Veerkamp et al., 1994; Veerkamp et al., 1995). Kolver et al. (2002) did report a significant genotype × diet interaction, when genotype was defined as New Zealand or overseas Holstein Friesian and diet was pasture based or total mixed ration.

Extensive research based on estimation of variance components has shown that genetic variance of milk production is environmentally sensitive if environments are defined in terms of for instance region (Carabaño et al., 1990; Ibáñez et al., 1999), weather information (Ravagnolo and Misztal, 2000), level of nutrition (Cromie, 1999), production level (De Veer and Van Vleck, 1987; Boldman and Freeman, 1990; Ibáñez et al., 1999; Berry et al., 2002; Calus et al., 2002; Kolmodin et al., 2002), herd size (Ibáñez et al., 1999), or other herd characteristics such as peak milk yield (Fikse et al., 2003a), or days open (Kolmodin et al., 2002). All these sources reported large heterogeneity of variances and some heterogeneity of heritabilities, but limited reranking of animals as the genetic correlations or correlations between breeding values for production traits in different environments were in most situations greater than 0.85.

When environments are defined as average production level, linear covariance functions have been fitted for yield traits (Calus et al., 2002; Kolmodin et al., 2002; Hayes et al., 2003), but higher order relationships have significantly been estimated (Veerkamp and Goddard, 1998).

Fertility, health and functional traits

Little research has been done on G×E in fertility, health and functional traits in dairy cattle. Pryce et al. (1999) found no genotype × feeding system interactions for health and fertility traits in dairy cattle. Estimated genetic correlations between SCS expressed in herd environments with low versus high average SCC were mainly close to unity (Castillo-Juarez et al., 2000; Raffrenato et al., 2003), apart from a reported value of 0.80 estimated for Swedish Holstein (Carlén et al., 2005), and as low as 0.83 when environments were defined based on management practices that enhance milk production (Raffrenato et al., 2003).

Covariance functions have been applied for fertility traits in a few studies, which reported heterogeneous heritabilities across fertility environments (Distl, 2001; Kolmodin et al., 2002). Reranking of animals was however limited, as reported genetic correlations were all close to unity.

The estimated genetic correlation for productive life of Swedish Red and White dairy cattle in contrasting herd environments was 0.74 (Petersson et al., 2005). Estimated genetic correlations for productive life between contrasting environments defined otherwise, were all closer to unity (Petersson et al., 2005).

Fatehi et al. (2003) reported genetic correlations less than unity for feet and leg characteristics between tie stall and free-stall ($r_g \geq 0.87$), free-stalls with solid or slatted floors ($r_g \geq 0.78$) or animals with intact or recently trimmed hoofs ($r_g \geq 0.88$).

In conclusion, despite the limited research on G×E in fertility, health and functional traits in dairy cattle, evidence has been found for the existence of G×E effects that are, in terms of genetic correlations across environments, of greater magnitude than yield traits.

Genetic correlations between traits

Milk yield has always been the major component of the breeding goal for dairy cattle, as milk production has a direct impact on the income of dairy farmers. However, selection focused mainly on milk yield may have led to an increase in the risk for some health and fertility disorders (Emanuelson, 1988; Pryce et al., 1998; Rauw et al., 1998). In 1994 only the Scandinavian countries considered fertility, calving performance, and health traits in their total merit index together with production traits (Philipsson et al., 1994). Nowadays, almost all countries participating in Interbull include those traits in their total merit index to increase production, fertility and health simultaneously. Philipsson et al. (1994) showed, based on a restricted total merit index, that the correlated responses in mastitis and fertility could be reduced to zero at the expense of only 12 to 15% of genetic gain in production.

In order to include multiple traits in a breeding goal, the genetic correlation between the traits need to be estimated. In the presence of G×E, genetic correlations between traits can change across environments. Although differences in genetic correlations estimated in different environments might partly result from estimation errors, the existence of differences across environments could be supported by the fact that genetic correlations between traits are sometimes different when estimated accurately in different countries. Castillo-Juarez et al. (2000) reported small changes in genetic correlations between mature equivalent milk yield, lactation mean SCS and conception rate at first service, across different production environments. Raffrenato et al. (2003) reported that correlations between milk yield traits and average somatic cell score in contrasting herd environments ranged from slightly favorable to antagonistic.

One way to allow genetic correlations between traits to vary across environments is application of a multitrait reaction norm model. Veerkamp and Goddard (1998) fitted a multitrait covariance function to allow the genetic correlations between milk, fat, and protein yield to change across stage of lactation and production level. Correlations between those traits expressed at different herd production levels ranged from 0.79 to 0.97. Kolmodin et al.

(2002) applied a multitrait random regression model for protein yield and days open and reported small changes in the genetic correlation between those traits expressed at different herd production levels.

Despite the limited research, some evidence has been found for environmental dependent genetic correlations in dairy cattle. The consequence is that changes in environment, i.e., management decisions, can possibly influence the negative relation between yield and health and fertility traits.

IMPORTANCE OF G×E
Selection, breeding programs & breeding goals

In a situation where G×E is present, it is important to ensure that the environment in which animals are selected, resembles the environment where the products of selection will be producing in in the future (Meuwissen, 1990; Visscher and Hill, 1992). For instance, a nucleus herd should resemble a commercial dairy herd in the future. Otherwise, the selection response in the production environment is a correlated response (**CR**), which is defined as $CR_Y = r_{XY}*(\sigma_Y/\sigma_X)*R_X$ (Falconer, 1952), where CR is the correlated response in environment Y, r_{XY} is the genetic correlation between the trait expressed in environment X and Y, σ_X and σ_Y are genetic standard deviations in environments X and Y, and R_X is the response of selection in environment X. If the genetic correlation (r_{XY}) between the trait expressed in environments X and Y is unity, i.e. there might be only scaling but no reranking of animals across environments, the CR still can change across environments, due to differences in genetic variances across environments. If the genetic correlation (r_{xy}) between the environments is less than 1, based on whether this correlation is smaller or larger than the break-even genetic correlation, different breeding programs for each environment can be considered or not (Mulder and Bijma, 2005).

Breeding organizations tend to focus on the 'average environment' rather than trying to address multiple breeding goals defined for different herd environments. No evidence is available that the 'national' breeding goal should be redefined in several smaller breeding goals, but it is known that dairy farmers with different farming styles tend to use different criteria when selecting bulls, probably based on their own philosophy and specific herd characteristics (Groen et al., 1993). The question is whether different farmers should use different bulls to optimize farm income, and if that is the case, whether application of one general breeding program leads to a situation where 'optimal' bulls are available for all farmers. Due to environmental sensitivity the 'average' animal might encounter serious problems in an extreme environment.

One way to prevent animals to be at risk in extreme environments is to select for environmental insensitivity in certain traits. However, those animals might perform not optimal in an environment in the other extreme of the environmental scale. Therefore, no

consensus exists whether selection should be for environmental sensitive or insensitive animals, also known as 'specialists' or 'generalists' (Fikse, 2002).

Socially acceptable animal production

Animal production is becoming more and more an activity in which a lot of different stakeholders are involved and demands of the consumers and the public opinion are becoming more important (Bennett, 1997). This puts several restrictions on animal husbandry. Food products should be safe, animal production should not have negative influences on the environment and at the same time a reasonable level of animal welfare has to be guaranteed.

Food safety, nitrogen management and animal welfare are becoming increasingly important issues in Europe. Although animal welfare is defined in several different ways, it is clear that animals should at least be able to function normally and be in good health (Sandoe et al., 1999). The public opinion that animals "should not suffer" can be translated into a minimum baseline with regard to a trait interfering with animal welfare. It is not the average achieved level of welfare that is most important, but likely the proportion of all animals that are below this baseline (Sandoe, 2004).

Different stakeholders have different interests and therefore an ongoing debate discusses whether the profit in breeding goals should be defined from the perspective of the farmer, the industry or the consumer (Goddard, 1998). So far, the farmer and the industry had most influence on the direction of breeding goal and selection, but the role of the consumer is becoming increasingly important, for instance due to recent trends such as socially acceptable animal husbandry (i.e. "Corporate Social Responsibility") or in some situations through governmental policy reflecting the opinion of consumers or even the community at large. Breeding goals could shift towards "maximization of what contributes to the quality of life of both production animals and humans" (Sandoe et al., 1999). If breeding goals are already defined in that way, the perception of "quality of life" is likely to change in time, implying that breeding goals will be continuously changing as well. Maximizing quality of life could result in a number of restrictions in practical circumstances, but might be benificial to the farmers as well, for instance if production increases due to better animal welfare. Although these restrictions could lead to changes in the production environment, farmers will still try to maximize their profits in the new situation and might need a different kind of animal to do so. Therefore, preservation of biodiversity can be regarded as a way to enable "maximization of quality of life" in the long term.

Socio-economic perspectives

The functioning of animals can be influenced both by changing genotype and environment. Changing the production environment is probably the best solution from the point of view of the public, because changing the animal, especially to enable animals to deal with their

production environment, could somehow interfere with the integrity of the animal (Sandoe et al., 1999). However, changes in the environment might not be acceptable, when these have to "correct" problems that are caused by modifying the animal through selection. The decision to change either the production environment or the genotypes will heavily depend on the costs of either change. For instance, if changing the production environment involves redesigning the interior of barns, the changes might be postponed until a new barn is built. Comparably, replacing all genotypes at once by for instance a different breed might be too expensive and the change of the genotypes might be done more gradually by changing the breeding goal and selection of other animals to breed the next generation.

Another motivation to use certain genotypes for animal production might just be the philosophy of the farmer or the farming system in which the farmer is producing (e.g., organic animal production). Different management styles have been identified, ranging from minimizing the production costs to maximizing the milk production per cow or from considering the cows to be the most important production factor to the machinery being most important (Groen et al., 1993; Dockes and Kling-Eveillard, 2004). Such philosophies might be an attempt to optimize the results in terms of for instance animal welfare, nutrient efficiency or simply the income of the farmer. If farmers do succeed to optimize their 'results' by choosing the 'optimal' genotype for their specific herd environment according to their own philosophy, then this would indicate the presence of G×E. Scientifically this kind of G×E might be hard to estimate, as comparison of different herds with different strategies might be complex, because several other things might be different across herds. Ideally, herds which have changed their genotypes throughout time should be used for such analysis, as time is needed for genotypes to 'adapt' (in this case in the meaning of 'being bred') to fit into their environments.

Biodiversity

The main focus of selection in production animals in the last century has been on production in industrial agriculture, characterized by its high input-output environment (FAO, 1998). This resulted in the use of a few highly productive breeds and a reduction of variation within those breeds (FAO, 1998). In a few simulation studies (Kolmodin et al., 2003; Van der Waaij, 2004) as well as in experiments (Falconer, 1990), it is shown that selection of environmentally sensitive animals based on phenotypic performance in a continuously improving environment, is expected to lead to increased environmental sensitivity of the selected animals, in the presence of G×E. This implies, that sudden changes in the environment, especially in the opposite direction of the environmental changes that occur during selection (e.g., changing back to a low input strategy or the outbreak of a disease), might lead to a rapid reduction of animal production and/or poor animal welfare.

Replacement of locally adapted breeds by exotic breeds can actually lead to a decrease in overall production, if the local circumstances are not sufficiently taken into account in the breeding goal (Sölkner et al., 1998). Also, use of locally adapted breeds might be more risk averse by ensuring a more or less stable production, which is quite important in marginal regions (Sölkner et al., 1998). This emphasizes the importance of maintaining biodiversity to off-set possible risks in different environments in the future. One of the main reasons to conserve a certain breed or strain is its specific adaptation to a certain environment, which is straightforward if conservation is focused towards a single environment. However, if conservation needs to address multiple environments, or rapidly changing environments, another approach might be to (partly) conserve breeds or strains that are relatively environmental insensitive and thus fit well in a range of environments.

CONCLUSIVE SUMMARY

Genotype × environment interaction is the phenomenon that genotypes express differently in different environments. This has basically the following consequences: 1) for different specific production environments, different optimal producing genotypes can be chosen or bred, and 2) environmental changes can lead to different phenotypic responses of different genotypes. In dairy cattle breeding programs, G×E for yield has been considered to be of minor importance, but is accounted for in international breeding value estimation. Little is however known on G×E for health and fertility traits. As estimates of G×E for health and fertility traits become available, the importance of G×E in breeding programs and goals might increase.

Chapter 3

Estimation of environmental sensitivity of genetic merit for milk production traits using a random regression model

M. P. L. Calus

R. F. Veerkamp

Animal Sciences Group
Division Animal Resources Development
P.O. Box 65, 8200 AB Lelystad, The Netherlands

ABSTRACT

The objective of this study was to estimate effects of environmental sensitivity of milk production traits for several environmental parameters and to investigate the impact of combining traits with different environmental sensitivity in an economic index. Variance components and breeding values were estimated for milk, fat, and protein yield and fat and protein percentage by applying a random regression on values of an environmental parameter for each sire. Fourteen environmental parameters were defined and fitted to data consisting of 151,696 heifers in 6780 herds in the Netherlands with first lactation records for milk production, somatic cell count, body condition score and number of inseminations. Milk, fat, and protein yield showed environmental sensitivity in combination with 12 environmental parameters. Herd-year averages of protein, body condition score, age at calving, calving interval and peak date of calving explained most genotype by environment interaction, mainly resulting from scaling effects. Almost all genetic correlations across environments were 0.99 or higher. Although heterogeneity of genetic variances was considerable, heterogeneity of heritabilities was limited. Scaling had a large effect on the weights of the economic index, but environmental sensitivities of milk, fat, and protein yields were approximately of equal magnitude. Consequently, very little reranking occurred based on the economic index.

INTRODUCTION

The phenomenon that different genotypes respond differently to changes in their environments is known as genotype × environment interaction (**G×E**) or as differences in environmental sensitivity (**ES**) of genotypes (Falconer and Mackay, 1996). This interaction can cause reranking of animals across environments or a change of scale, i.e., variance, across environments (Lynch and Walsh, 1998). Traditionally, animal breeders are more concerned about reranking of animals than about differences in scale between environments because differences in scale do not affect the ranking of the animals for the considered trait. However, reranking of animals across environments is limited for milk production traits (Veerkamp et al., 1995; Cromie, 1999; Calus et al., 2002), although there is evidence that variances and heritabilities vary (Veerkamp et al., 1995; Cromie, 1999; Calus et al., 2002). Scaling effects can be accounted for in the breeding value estimation model (Meuwissen et al., 1996) and do not influence the ranking of sires, based on a single trait. However, if scaling effects are different for traits that are combined together in an economic index, the relative importance among the traits might change and cause reranking based on this economic index (Namkoong, 1985). In that case it might be more appropriate to include environmental sensitivity in breeding decisions rather than correct for it in the statistical model.

An environmental parameter (**EP**) reflects the environment encountered by the animals. An EP can reflect production level of a herd (Veerkamp and Goddard, 1998; Calus et al., 2002; Kolmodin et al., 2002), or other characteristics of the herd, such as for instance average

calving interval or average age at calving (Fikse et al., 2003a). Describing G×E for dairy cattle with a covariance function of an EP is recently described for a limited number of EP (Veerkamp and Goddard, 1998; Calus et al., 2002; Kolmodin et al., 2002; Fikse et al., 2003a). The use of an EP in a covariance function has the advantage that environments are treated as a continuum, rather than a set of arbitrarily defined groups of the data. The EBV of an animal, which is divided into an environment independent and an environment dependent part, is also called reaction norm (Falconer and Mackay, 1996). The simplest form of a covariance function describes the ES of the genotype as a linear function of the EP, but higher order functions are possible. Covariance functions can be estimated by a two step procedure or by random regression models (Van der Werf et al., 1998).

The objective of this paper is to estimate ES for milk production traits for a range of EP in order to identify those EP that gave most ES and to investigate the effects of ES on reranking in the Dutch economic index (INET) combining milk, fat, and protein yields.

MATERIALS AND METHODS
Data

The data contained first-lactation test-day records for milk production traits and SCC, date of first calving, date of each insemination and body condition scores for 271,606 heifers calving in 1998 and 1999 from 12,347 dairy herds in The Netherlands. A few criteria were used to select the data for the estimation of variance components of milk production traits. Each heifer needed at least five test-day records during the period of 5 to 305 DIM of which at least one was on or after 180 DIM; heifers that calved at an age of less than 640 or more than 1310 d were deleted. These criteria decreased the number to 267,120 heifers in 11,602 herds. Selected animals were deleted if their herd did not meet the criteria of the EP, which are explained later. To calculate EP, for each herd-year a minimum number of four records was needed for calving interval, as this EP was most restrictive, and five records for other EP. Effectively, a total of 151,696 mostly Holstein-Friesian and Meuse-Rhine-Yssel heifers in 6780 relatively large herds were selected. Each herd-year contained an average of 14 heifers.

In the pedigree file, a maximum of five generations of sires were included, together with the pedigrees of dams of first and second generations of sires and the sires' maternal grandsires. A total of 4769 sires with daughters in the data were identified. Sires had on average 32 daughters. The relationship matrix contained 14,382 animals.

Traits

Five traits were evaluated: average daily milk yield, fat yield, protein yield, fat percentage, and protein percentage. Milk, fat, and protein yield were calculated as the average of the test-day yields between 5 and 305 DIM. Fat and protein percentage were calculated as the average yield of fat and protein divided by the average milk yield.

Environmental parameters

Environmental parameters were calculated for each herd-year level, based on calving date. Potentially a large number of EP could be defined, but parameters used here were chosen because they: 1) reflect management and environment, 2) are obtainable from the available data, 3) are continuous rather than categorical, i.e., the parameter is expressed on a scale rather than defined in several classes, and 4) are not too strongly correlated with each other. Each parameter was averaged over all heifers calving in the relevant herd-year. For parameters that reflected traits with more records per lactation, first an average was calculated for each selected animal.

Test-day record parameters. Average protein and SCC were calculated from test-day records in 1998 and 1999. Each SCC test-day record was transformed to a SCS, by SCS = $\log_{10}$(SCC). Persistency was calculated in two different ways. First, persistency was calculated for each animal from the ratio of milk production on the test-day closest to 60 DIM to milk production on the test-day closest to 240 DIM (Zwald et al., 2001), both in a range of 42 d around those. Second, persistency was taken as the highest test-day milk production of a heifer divided by its average test-day milk production. This last parameter is called relative peak milk yield.

Age at calving and herd size parameters. Average age at calving and number of freshened heifers were calculated over all heifers that calved during the year, regardless how many days they produced. In both years, the change in number of freshened heifers was set to the difference between 1998 and 1999.

Energy balance parameters. Energy balance reflects the ability of management to tune the feed intake to the energy requirements and therefore indicates whether tissue reserves are mobilized or deposited in the cow. Body condition score reflects cumulated energy balance (Chilliard et al., 1991). Body condition score was measured during classification and only once during the first lactation. Average body condition score (**BCS**) was calculated from all classified heifers in a herd that calved in the same year. Other traits that reflect energy balance are change in fat percentage and fat over protein ratio (De Vries and Veerkamp, 2000). Change in fat percentage was calculated as the difference in fat percentage on the test-day closest to 77 DIM and the test-day closest to 14 DIM (De Vries and Veerkamp, 2000), both in the range from 10 to 100 DIM. Fat over protein ratio was calculated by test-day and then averaged across test-days.

Calving and insemination parameters. Calving interval is the period between first and second calving. The number of inseminations required for a successful second calving was estimated during first lactation. The herd calving pattern was represented by peak date of calving and distribution of calving dates over the year. The peak date of calving shows the date around which the heifers are calving and the distribution of the calving dates shows whether the heifers are calving near that date or throughout the whole year. Average day of

calving can be calculated by numbering each day of the year from 1 to 365 and obtaining the average of the renumbered calving dates (Zwald et al., 2001). A disadvantage of this method is that a non-equal distribution of calving during a year, e.g., if the peak date of calving does not fall on July 1st, leads to an underestimation or an overestimation of the peak date of calving. Here, a slightly different procedure is used. The peak date of calving is calculated by iteratively repeating the following procedure: 1) calculate the average of calving dates (Zwald et al., 2001) and 2) define a maximal time period in the same calendar year with the average of step 1 as central point. For instance, with an average calving date of d 140, the new period ranges from d 0 to 280. If the period became shorter than 182 d it was expanded over the borders of the calendar year. Insemination data from 1997 and 2000 were available to make this expansion feasible. The time period in step 1 is in the first iteration the calendar year and in later iterations the period defined in step 2 in the previous iteration. Step 1 and 2 are repeated until the average day of calving no longer changes. The converged average day of calving is considered to be the peak date of calving. For a few herds, two peaks of calving in a year made convergence impossible. If convergence had not occurred after 1000 iterations, the average day of calving of all 1000 iterations was taken. Values for 1999 were adjusted by substracting 365 days to come to the same standard as in 1998. The distribution of calving dates during the year was calculated by the use of an interval of 182 d and one of 365 d, both centered on the peak date of calving. The distribution of calving dates was calculated as the ratio of the number of calving dates in the short interval to the number of calving dates in the long interval.

Estimating variance components and environmental sensitivity

Variance components were estimated by using a sire model. Environmental sensitivity was modeled by applying a random regression for each sire, representing its EBV, on values of an EP for the herds in which his daughters were producing. A fixed linear and quadratic regression for age at calving was included, as was a fixed effect to account for herd-year-season (**HYS**) groups. Furthermore, a fixed polynomial was also applied to the EP, to account for the average effect in each environment. An HYS effect was not fully covered by the herd-year effect. In each situation, only one EP was used for both the random regression for sires and the fixed regression for herd-year. The residual variance was calculated for 10 equally sized groups, based on increasing EP, to include heterogeneous residual variances in the model.

The HYS groups were defined by a method that optimizes the composition of HYS groups based on the calving dates and intervals between consecutive calving dates in a herd (Crump et al., 1997). Initially, the criteria of a maximum period of 91 d and a minimum of five animals per HYS group were applied. If some animals were not assigned to an HYS group based on these criteria, they were forced to join one by relaxing the criterion for the maximum

period. The same was applied to animals from groups that had fewer than five animals, to force them to join another HYS group.

The applied model was:

$$Y_{ijklmnoq} = \mu + HYS_i + \gamma_0 * AGE_j + \gamma_1 * (AGE_j)^2 + \sum_{k=0}^{10} \beta_k P_{km} + \sum_{l=0}^{S} \alpha_{ln} P_{lm} + E_{ijklmnoq}$$

where $Y_{ijklmnoq}$ is the performance of heifer q; μ is the average performance over all animals, HYS_i is the effect of herd-year-season group i; γ_0 and γ_1 are coefficients of linear and quadratic fixed regression on age at calving j in days, respectively; AGE_j is age at calving in days of heifer q; β_k is coefficient k of a fixed regression on element k of the orthogonal polynomials of all environments; P_{km} is element k of the orthogonal polynomial resembling an environmental parameter of environment m; α_{ln} is coefficient l of the random regression on the orthogonal polynomials of all environments of the daughters for sire n; P_{lm} is element l of the orthogonal polynomial resembling an environmental parameter of environment m; s is the largest significant coefficient l of the random regression; and $E_{ijklmnoq}$ is the residual effect of heifer q in environment m within group of environments o ($o = 1, 2, ..., 10$).

The order of the polynomials for the fixed regression on an EP was arbitrarily set to 10 in each situation. For the random regression, the order of the polynomial was increased per combination of trait and EP until the extra added components of the next order did not significantly improve the fit of the model or the variance of the extra component was zero. The log likelihood ratio test (Kirkpatrick et al., 1990) was used to compare the fit of two models with consecutive orders of polynomials.

The sire variances for values of an EP are calculated as $\Phi S \Phi'$, where Φ is a matrix with polynomial coefficients for a value of the EP on each row and S is an $n \times n$ matrix, where n is the highest order of the polynomial + 1, with variances of each random regression coefficient on the diagonal and covariances between the random regression coefficients on the off-diagonals. The residual variance was calculated for 10 different groups. Residual covariances between groups were assumed to be zero. Covariances between sire and residual effects were assumed to be zero and not taken into account.

Several criteria can be defined to rank EP based on the given amount of G×E. In this study we used the absolute change in sire variances between the 25 and 75% of the environmental scale as an indicator of change in sire variance across environments and therefore of the given amount of G×E. To check the results of the random regression model, a multitrait model was applied for selected combinations of traits and EP. The program ASREML (Gilmour et al., 2002a) was used for all analyses.

Economic index

The Dutch INET is an economic index that includes milk, fat, and protein yield and is calculated as INET = -0.08 × EBV(milk yield) + 1 × EBV(fat yield) + 6 × EBV(protein

yield). The INET was used to investigate the effects of ES on the combination of milk, fat, and protein yields. First, the INET was calculated based on the results of the described model. No base adjustments in the index were made, i.e., average breeding values were not adjusted based on the average of the whole current population. Secondly, the economic weight was readjusted to real economic weights in a few environments. The correlated response of selection in a different environment is:

$$CR_Y = r_{XY} \times (\sigma_Y/\sigma_X) \times R_X \text{ (Falconer and Mackay, 1996)},$$

where CR is the correlated response in environment Y, r_{XY} is the genetic correlation between environment X and Y, σ_X and σ_Y are genetic standard deviations in environments X and Y, and R_X is the response of selection in environment X. The adjustment factor of the weights of the INET is then equal to $r_{XY}*(\sigma_Y/\sigma_X)$. If no reranking occurs or adjustment is for scaling effects only, i.e., $r_{XY} = 1$, this reduces to (σ_Y/σ_X).

RESULTS
Environmental parameters
Mean and range for the environmental parameters is given in Table 3.1 based on 151,696 heifers in 6780 herds. Herds differed considerably for the EP, for example, average protein ranged from 0.46 kg of protein per day to 1.21 kg of protein per day, average age at calving from 672 to 1028 d and herd size from 5 to 270 numbers of heifers calving. Correlations among all EP ranged from –0.43 to 0.45 but were generally weak. Strongest correlations were found for pairs of EP that were calculated from the same traits, such as EP defined on production traits and defined on the number of freshened heifers in a herd. As peak date of calving is one of the newly defined EP and had a considerable effect, the number of animals with peak date of calving in a certain month is shown for both years (Figure 3.1). This illustrates that most herds have their heifers calving in the autumn, although this is less clear for 1999 than for 1998.

Significant reaction norms were found in 50 out of 70 combinations of traits and EP. The highest estimable and significant order of the polynomial is given for each combination of traits and EP in Table 3.2. Milk, fat, and protein yield each had significant reaction norms in combination with 12 EP. These three traits generally showed the same order of significant polynomials for a given EP. Fat and protein percentage had significant reaction norms in combination with 6 and 8 EP.

Table 3.1. *Mean, standard deviation, coefficient of variation and range of all environmental parameters.*

Environmental parameter	Mean	SD	CV (%)	Minimum	Maximum
Average protein (kg/d)	0.84	0.09	10.7	0.46	1.21
Fat/protein	1.26	0.05	4.1	1.02	1.51
SCS	2.55	0.99	38.8	-5.74	5.54
Persistency	1.10	0.04	3.6	0.99	1.47
Relative peak milk yield	1.23	0.05	3.7	1.09	1.57
Age at calving (d)	797	39	4.9	672	1028
Number of animals	22.73	16.22	71.3	5	270
Change in number of animals[1]	3.95	9.35	237	-40	71
BCS (thin-fat: scale 1-9)	4.75	0.66	13.8	2.18	7.50
Change in fat percentage (%)[2]	-0.20	0.26	-125.6	-1.53	1.33
Calving interval (d)	388	27.1	7.0	316	605
Number of inseminations	1.33	0.25	19.0	1.00	3.50
Peak calving date (d)[3]	232	78	22.4	-108	434
Distribution of calving dates	0.41	0.18	44.9	0.00	1.00

[1]In both years the difference between 1998 and 1999 is used.

[2]The difference in fat percentage on the test day closest to 77 DIM and the test day closest to 14 DIM.

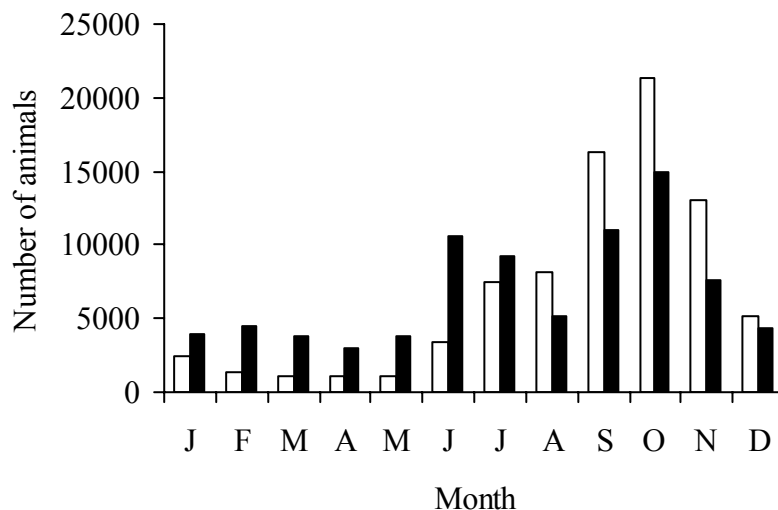[3]-108 means day 257 of the year before, 434 means day 69 of the next year.

Figure 3.1. *Distribution of the numbers of animals with a peak date of calving in a certain month for both 1998 (□) and 1999 (■).*

Table 3.2. *The number of the highest estimable and significant (P < 0.05) polynomial order[1] of the combination of the environmental parameter in the row and the trait in the column.*

Environmental parameter	Traits				
	Milk	Fat	Protein	Fat %	Protein %
Average protein	2	2	2	2	2
Fat/protein	2	2	3	2	0
SCS	2	2	2	0	1
Persistency	2	2	2	0	0
Relative peak milk yield	1	2	1	0	2
Age at calving	1	1	1	1	2
Number of animals	2	1	2	1	2
Change number of animals	2	2	2	0	0
BCS	2	2	2	1	1
Change in fat percentage	0	0	0	0	1
Calving interval	2	1	1	1	0
Number of inseminations	1	2	1	0	0
Peak calving date	3	2	1	1	0
Distribution of calving dates	0	0	0	0	1

[1]A zero means that the first order polynomial was not significant.

Environmental sensitivity

The EP for the environments at 25, 50 and 75% of the data based on the increasing environmental scale were used to calculate sire variances and heritabilities for these environments for milk, fat, and protein yield. The results for the sire variances calculated for 305 days are shown in Table 3.3. The heritabilities were comparable across environments for all combinations of traits and EP.

Environmental parameters were ranked on decreasing absolute change in sire variances between 25 and 75% on the scale of the EP, indicating decreasing G×E for milk, fat, and protein yield (Table 3.4). For the trait protein yield, average protein, BCS, calving interval, age at calving and persistency were the most important EP (Figure 3.2). Age at calving and calving interval showed the same pattern of sire variance, which decreased by a third across the environmental scale. The sire variance showed the most curvilinear relationship with average protein. For average protein and BCS, the sire variances doubled across the environmental scale. For the EP average BCS, the sire variances for milk, fat, and protein yield are shown in Figure 3.3. The environmental scale in Figure 3.3 resembles the interval of the mean EP ± two standard deviations. The sire variances are following a similar pattern across the environmental scale.

Table 3.3. *Sire variances in kg^2 per 305 d at 25, 50 and 75% on the scale of the environmental parameters, for all combinations of environmental parameters and the traits milk, fat, and protein yield that showed a significant environmental sensitivity.*

Environmental parameter	Milk			Fat			Protein		
	25%	50%	75%	25%	50%	75%	25%	50%	75%
Average protein	116,598	131,872	142,291	169.9	189.8	203.8	92.4	107.8	119.5
Fat/protein	118,942	118,923	118,039	168.7	170.5	171.8	96.9	95.9	94.3
SCS	116,095	114,979	117,118	165.9	169.0	175.8	91.9	91.6	95.1
Persistency	114,532	118,291	122,291	164.4	169.8	175.5	91.7	94.9	99.0
Relative peak milk yield	118,439	117,388	116,625	172.9	172.6	170.9	94.2	94.3	95.2
Age at calving	123,295	118,942	114,049	177.0	171.5	165.5	103.1	97.1	90.2
Number of animals	117,202	116,905	116,942	165.6	166.8	169.0	93.3	93.4	93.8
Change nr. of animals	119,379	117,063	114,644	166.6	167.8	169.8	96.1	93.8	91.7
BCS	109,639	120,505	127,984	154.3	174.8	189.2	85.4	97.3	106.1
Calving interval	124,533	115,965	109,100	176.7	170.7	164.2	102.7	96.3	89.4
Nr. of inseminations	115,072	116,653	119,537	166.7	168.6	172.3	92.0	93.8	97.0
Peak calving date	112,881	118,264	120,046	154.7	164.6	172.1	89.5	94.0	96.2

Table 3.4. *Environmental parameters ranked per trait for decreasing absolute change of the sire variances between 25 and 75% on the environmental scale.*

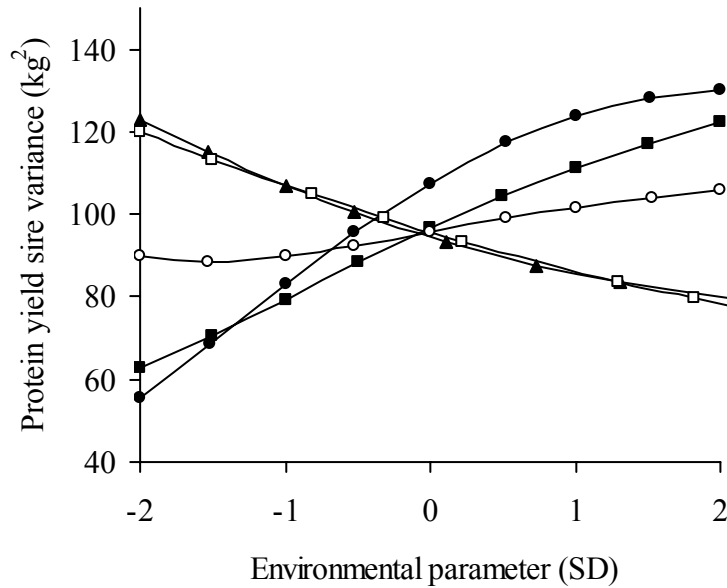| Environmental parameter	Traits		
	Milk	Fat	Protein
Average protein	1	2	1
Fat/protein	11	11	10
SCS	10	7	9
Persistency	5	6	5
Relative peak milk yield	9	12	11
Age at calving	4	5	4
Number of animals	12	9	12
Change number of animals	7	10	8
BCS	2	1	2
Calving interval	3	4	3
Number of inseminations	8	8	7
Peak calving date	6	3	6

Figure 3.2. *Sire variance of protein yield in kg² for average protein (protein) (●), BCS (■), calving interval (▲), age at calving (□), and persistency (○) given the deviation of the mean environmental parameter in standard deviations.*

The genetic correlation between a trait in the environment at 25% and the same trait at 75% of the data on the environmental scale were calculated for the model with the highest significant order for the given combination of trait and EP. The combination of the trait protein and the EP fat over protein ratio gave a genetic correlation of 0.96 (SE = 0.009). All other correlations were 0.99 or higher (SE ranging from 0.000 to 0.004). The multitrait model (results not shown) gave comparable results to those of the random regression model.

Economic index

The overall economic value for milk, fat, and protein yield, called INET, was calculated on the scale of the EP average BCS for the 10 sires with the highest number of daughters in the data (Figure 3.4). These bulls are not representative for all bulls in the population, but they represent the breeding bulls that are widely used by dairy farmers. The INET increases with increasing average BCS and little reranking happened. The sire with the highest change of INET shows an INET of 65 Euros for herds with an average BCS of 3.5 and an INET of 120 Euros for herds with an average BCS of 6.0.

Based on the scaling effects of milk, fat, and protein yield across environments, the economic weights in the INET formula were adjusted. The results for the EP average BCS are shown in Table 3.5. The change of the economic weights with the change of the EP clearly illustrates the effect of scaling.
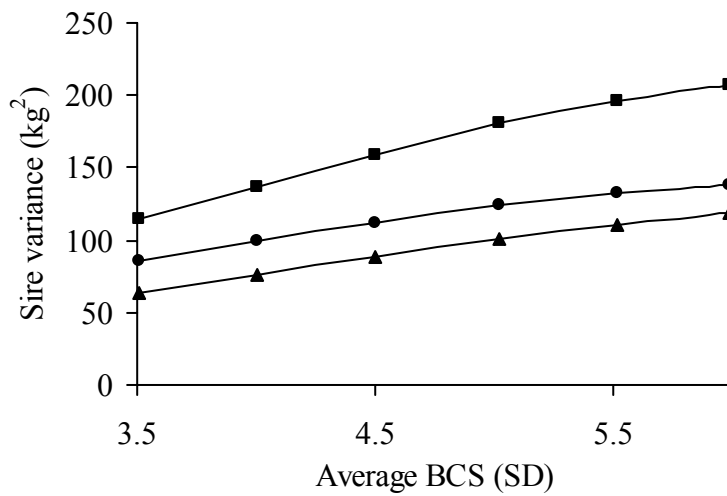
Figure 3.3. *Sire variances for 305 d milk/1000 (●), fat (■), and protein yield (▲) as a function of the average BCS per herd-year.*
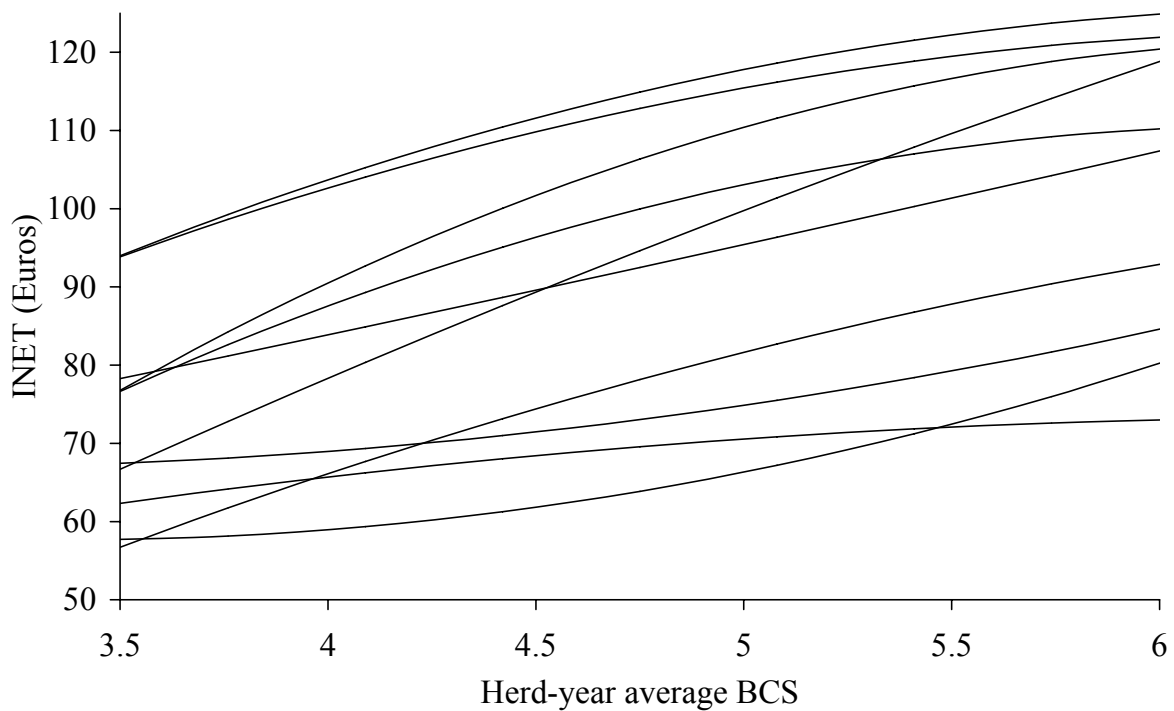

Figure 3.4. *The Dutch total economic value for milk, fat, and protein yield (INET) in Euros, of the 10 sires with the highest number of daughters in the data, as function of the average BCS per herd-year.*

Table 3.5. *Economic values (in Euros) for 305-d milk, fat, and protein yield in the INET formula, for different values of the environmental parameter average BCS.*

BCS	Milk	Fat	Protein
3.5	-0.068	0.82	4.90
4.0	-0.074	0.90	5.36
4.5	-0.078	0.97	5.79
5.0	-0.082	1.03	6.19
5.5	-0.084	1.07	6.49
6.0	-0.087	1.10	6.74

DISCUSSION

Environmental parameters

Herds were required to have at least four records for calving interval and at least five records for all other EP in order to be selected. This criterion led to a substantial loss of data. However, it ensured that small herds with very few informative animals were not taken into account. The data from such herds might be biased by the small numbers and from the fact that their genetic composition is more likely to be nonrandom. The edited data still contained herds that differed considerably for EP, indicating that data editing did not discard certain types of herds, apart from small ones.

Twelve EP gave significant reaction norms for milk, fat, and protein yield. Results for six of these parameters for milk yield were reported in literature. Significant ES for the EP herd size, relative peak milk yield, persistency and age at first calving were found (Fikse et al., 2003a), but no significant reaction norms for the EP fat over protein ratio and calving interval were found in that study. Average protein and BCS showed the highest ES based on our definition. Average protein is proven to be a useful EP before (Calus et al., 2002; Kolmodin et al., 2002). These results indicate that adjusting for heterogeneous variances in the genetic evaluation model might be possible by using EP.

The situation in which the EP is calculated based on the evaluated trait, needs some attention. The fact that EP are calculated as phenotypic averages within a herd implies that the breeding values are partly based on the breeding values of sires of the cows producing in a herd. Kolmodin et al. (2002) suggested that this may not be a problem, but no extensive evidence was given. In the case of random use of sires across environments, this will probably not be a problem, as the average breeding value will be zero, but the precise implication of including the evaluated trait in the calculation of the EP is not clear yet.

It is tempting to suggest that significant scaling effects shown for some of the EP, might be a result of the association with yield, because generally an increasing level of production also leads to increasing variance. In this study both BCS and calving interval showed changing sire variances with increasing values of the EP (Table 3.3), and peak date of calving and BCS

had correlations of 0.19 and -0.01 with average protein. However, herds with an average BCS of four and herds with an average BCS of six, both had an average 305-d protein production of 250 kg, indicating that only limited effect can be expected from the low correlation between protein yield and BCS. Still, it illustrates the problem of interpreting EP, and the need to consider multiple EP simultaneously.

Body condition score gave considerable scaling effects. Body condition score reflects cumulated energy balance (Chilliard et al., 1991). It was reported (Cromie, 1999) that defining environments based on the amount of concentrate fed causes scaling effects that are comparable to scaling effects if environments are defined on average protein. These results support the suggestion of others (Calus et al., 2002; Kolmodin et al., 2002) that feed intake and feed supply are important parameters in relation to environmental sensitivity of genetic merit for milk yield.

Environmental sensitivity

Environmental sensitivity was defined by Falconer (1990) as the difference between phenotypic values of a genotype or a population in two environments, divided by the difference of the means of all individuals in both environments. In our study, ES is defined at a population level as the variance in reaction norms of genotypes.

Genetic correlations of a trait across environments were high, indicating that reranking hardly occurred across environments, as was expected from literature (Calus et al., 2002; Kolmodin et al., 2002; Fikse et al., 2003a). However, sire variances showed considerable scaling effects for a number of the EP. At the same time, heritabilities were comparable across environments, indicating that scaling effects for environmental and sire variances were comparable. Heterogeneous heritabilities for comparable models were reported, but in these studies the heterogeneity of residual variances was not taken into account (Calus et al., 2002; Kolmodin et al., 2002). If heterogeneity of sire variances is accommodated in the model, but heterogeneity of residual variance is not, the presence of scaling effects is likely to cause heterogeneity of heritabilities.

Herds with high protein, high persistency, young age at calving, high BCS, short calving intervals and calving peak in the fall or winter appeared to have the highest genetic variance for milk, fat, and protein yield. This means that herds that have one or more of these characteristics are more likely to benefit more from the use of bulls with high genetic merit and the use of expensive high genetic merit bulls is more easily justified in those herds. At the same time, selection of animals on those herds will be more effective when there is an insufficient correction for heterogeneous variances in the breeding value estimation.

Economic index

The economic value of a trait was affected if the trait showed large scaling effects. If scaling effects are different among traits, the relative importance of these traits in an economic index can change (Namkoong, 1985). The economic index might give different selection responses depending on herd environment, and therefore reranking across environments might occur. This might reduce the total benefit of selection based on this economic index. In this study, the economic values of the traits in the index were only adjusted for scaling effects. Genetic correlations among environments of the adjusted economic values (Table 3.5) ranged from 0.93 to 0.99. Taking these into account would cause greater differences across environments. As shown here, reranking based on INET will be small, for a number of reasons. First, the scaling effects of milk, fat, and protein yield are comparable across environments. Second, the genetic correlations among these three traits are high and therefore the economic index is relatively insensitive for changes in economic values (Veerkamp et al., 1995). However, if other traits are included in the economic index, with scaling effects that are independent from those of production traits, the scaling effects could cause considerable reranking based on the economic index (Namkoong, 1985). This clearly indicates that scaling effects might be of importance in animal breeding programs.

CONCLUSIONS

Herds with high protein, high persistency, young age at calving, high BCS, short calving intervals, and calving peak in the fall or winter, have higher variances for the yield traits and are therefore expected to benefit more from the use of bulls with high genetic merit and selection of animals based on those herds will be more effective. Scaling effects for milk, fat, and protein yield were considerable, but comparable, indicating that no large differences of environmental sensitivities among these traits were found on a population level. Therefore, reranking based on economic index was limited. The absence of reranking based on a single trait does not necessarily mean that G×E is not important and scaling effects can easily be accounted for by adjusting the data. As more reproduction and health traits are included in total merit indices, further research is needed to explore the ES of these traits.

ACKNOWLEDGEMENTS

Chapter 4

Effects of data structure on the estimation of covariance functions to describe genotype by environment interactions in a reaction norm model

M. P. L. Calus[1]

P. Bijma[2]

R. F. Veerkamp[1]

[1]Animal Sciences Group
Division Animal Resources Development
P.O. Box 65, 8200 AB Lelystad, The Netherlands

[2]Animal Breeding and Genetics Group
Department of Animal Sciences
Wageningen University PO Box 338
6700 AH Wageningen, The Netherlands

ABSTRACT

Covariance functions have been proposed to predict breeding values and genetic (co)variances as a function of phenotypic within herd-year averages (environmental parameters) to include genotype by environment interaction. The objective of this paper was to investigate influence of definition of environmental parameters and non-random use of sires on expected breeding values and estimated genetic variances across environments. Breeding values were simulated as a linear function of simulated herd effects. The definition of environmental parameters hardly influenced results. In situations with random use of sires, estimated genetic correlations between the trait expressed in different environments were 0.93, 0.93 and 0.97 while simulated at 0.89 and estimated genetic variances deviated up to 30% from simulated values. Non random use of sires, poor genetic connectedness and small herd size had a large impact on estimated covariance functions, expected breeding values and calculated environmental parameters. Estimated genetic correlations between a trait expressed in different environments were biased upwards and breeding values were more biased when genetic connectedness became poorer and herd composition more diverse. The best possible solution at this stage is to use environmental parameters combining large numbers of animals per herd, while loosing some information on genotype by environment interaction in the data.

INTRODUCTION

The application of genetic covariance functions (CF), to model traits in dairy cattle by predicting breeding values as a function of an environmental parameter (EP), was suggested several times (Veerkamp and Goddard, 1998; Kolmodin et al., 2002; Calus and Veerkamp, 2003; Fikse et al., 2003a). The change of an animal's expected breeding value (EBV**)** across environments represents its environmental sensitivity. The CF includes differences in environmental sensitivity of genotypes for a trait, also known as the genotype by environment interaction (G × E), in the variance components, regardless whether it originates from scaling effects or re-ranking of animals across environments. This is in contrast to the usually applied methods for breeding value prediction that either (1) ignore environmental sensitivity or (2) ignore re-ranking by correcting only for heterogeneity of variances (Meuwissen et al., 1996). In international breeding value estimation, where G × E is included in the model by regarding records of animals in different countries as different traits (Schaeffer, 1994), both scaling and re-ranking are considered. However, this method has several limitations, for example the grouping of animals based on country borders while herd environments in small neighbouring countries may be much more similar than herd environments in different parts of a large country (Weigel and Rekaya, 2000). Also, a large number of countries implies a large number of traits, which increases the chance that the estimated genetic covariance matrix is not positive definite (Hill and Thompson, 1978), indicating that problems are likely to appear in the estimation of variance components for such multitrait models. Therefore, application of

CF is of interest to take G × E into account for example in international breeding value estimation, or to investigate importance of G × E.

In applications in dairy cattle, an EP is usually calculated as the mean phenotypic performance of a trait in an environment (Veerkamp and Goddard, 1998; Kolmodin et al., 2002; Calus and Veerkamp, 2003; Fikse et al., 2003a), which implies that both average genetic level within the herd and the animals own true breeding value (TBV) are included in the EP (Kolmodin et al., 2002; Calus and Veerkamp, 2003; Fikse et al., 2003a). Confounding between EP and TBV might affect EBV for example in herds with a non-average genetic composition or relatively small herds, since it might be difficult to disentangle genetic and environmental effects. Kolmodin et al. (2002) tried to partly solve this problem by calculating EP from more animals in the herd, rather than only from animals whose sires are being evaluated. Another problem with application of CF is that low numbers of daughters per sire might lead to problems in predicting breeding values. The number of records from daughters of a sire is the number of data points through which the curve representing the sires' EBV is fitted and extrapolation of curves of sires with a low number of daughters to extreme environments might be required. Another typical animal breeding problem is that herds with better management tend to use different sires than herds with a low level of management. This might lead to poorer genetic connectedness between herd environments but also to a covariance between genotype and herd environment. Hence, it is not known whether CF can handle these typical animal breeding problems, such as limited genetic connectedness between herds or preferential treatment, that exist in both within country and international breeding value estimation.

The objective of this paper was to investigate influence of definition of EP and levels of preferential sire use in herds on expected breeding values and estimated genetic variance across the range of EP in one population by stochastic simulation. Data structures were varied by changing the number of daughters per sire and average number of animals per herd for traits with low and high heritabilities, applying three levels of G × E interaction.

MATERIALS AND METHODS
Simulation

Data were simulated to compare estimated variance components, calculated EP and expected breeding values from different models. A record was simulated including the animals breeding value, a herd effect and a residual. A breeding value (a) was simulated as the average of the parents breeding values plus a Mendelian sampling term (ms). Each component included an intercept (a_0 and ms_0) and a linear regression on the environment (a_1 and ms_1):

$$a = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \frac{1}{2}\left(\begin{bmatrix} a_0 \\ a_1 \end{bmatrix}_{sire} + \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}_{dam} \right) + \begin{bmatrix} ms_0 \\ ms_1 \end{bmatrix},$$

where

$$Var(a) = \begin{bmatrix} \sigma^2_{a_0} & \sigma_{a_0,a_1} \\ \sigma_{a_0,a_1} & \sigma^2_{a_1} \end{bmatrix}, \quad ms = \begin{bmatrix} ms_0 \\ ms_1 \end{bmatrix} \sim N(0, Var(ms))$$

and

$$Var(ms) = \begin{bmatrix} \tfrac{1}{2}\sigma^2_{a_0} & \tfrac{1}{2}\sigma_{a_0,a_1} \\ \tfrac{1}{2}\sigma_{a_0,a_1} & \tfrac{1}{2}\sigma^2_{a_1} \end{bmatrix}.$$

The Mendelian Sampling term was simulated dependent on environment, to ensure that it explained half of the total genetic variance in each given environment. The breeding value in a specific environment with a simulated herd effect *herd* was calculated as: $TBV_{herd} = z_{herd}'a$,

where $z_{herd} = \begin{bmatrix} z_{0_{herd}} \\ z_{1_{herd}} \end{bmatrix}$, and $z_{0_{herd}}$ and $z_{1_{herd}}$ are respectively the level and slope of the animals

breeding value. TBV_{herd} had a normal distribution $N(0, z'_{herd}Var(a)z_{herd})$ for each value of *herd*. Application of reaction norm models as a function of herd average of the analysed trait showed that genetic variances increase with increasing herd level of the trait (Kolmodin et al., 2002; Calus and Veerkamp, 2003). In order to simulate mainly increasing genetic variance across environments, 99% of simulated herd effects (*herd)* got positive simulated values by sampling from a normal distribution $N(1, 1/9)$. The *residual* was simulated homogeneously across environments by sampling from a normal distribution $N(0, \sigma_e^2)$, where $\sigma_e^2 = 1 - \sigma^2_{a_0}$.

$\sigma^2_{a_0}$ and $\sigma^2_{a_1}$ were set to 0.04 and 0.02 to reflect a low heritability trait (e.g., a fertility trait) and to 0.4 and 0.2 to reflect a high heritability trait (e.g., a milk production trait). The correlation between level and slope (r_{a_0,a_1}) was set to –0.5, 0 or 0.5. The simulated genetic correlation between the trait expressed in different environments was calculated by dividing the genetic covariance between two environments, with simulated herd effects of $herd_1$ and $herd_2$, by the square root of the product of the genetic variances in both environments:

$$r_{g_{herd1herd2}} = \frac{z'_{herd_1}Var(a)z_{herd_2}}{\sqrt{z'_{herd_1}Var(a)z_{herd_1} * z'_{herd_2}Var(a)z_{herd_2}}} \qquad (1)$$

As a result of the chosen variances, both the low and high heritability traits had simulated values for $r_{g_{(herd=0.5,herd=1.5)}}$ of 0.74, 0.89 and 0.96 representing different amounts of re-ranking for r_{a_0,a_1} being respectively –0.5, 0 or 0.5. Simulated heritabilities across environments for both the low and high heritability traits are shown in Figure 4.1.
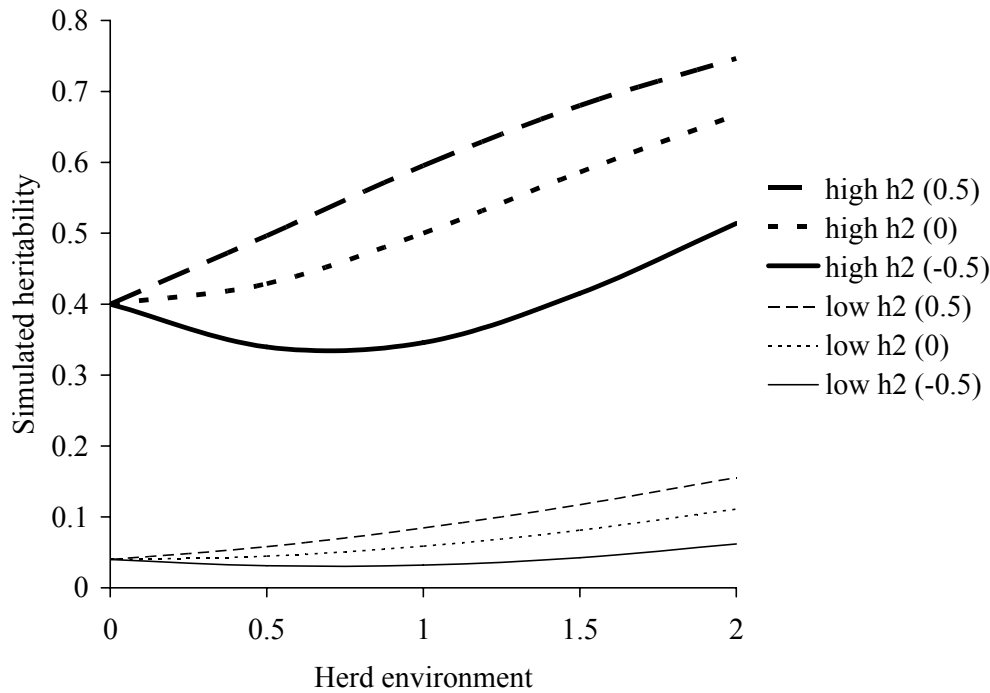
Figure 4.1. *Simulated heritabilities of the low and high heritability trait as function of the herd environment for situations with correlations between level and slope of –0.5, 0 and 0.5.*

Population structure

Different values were considered for the input parameters (Table 4.1). All values in bold were used as default in situations where different values were considered for the other parameters. A simulated population contained 50,000 animals, 500 or 2000 sires and 1000 or 5000 herds. The number of daughters per sire was 25 or 100. The average number of animals per herd was 10 or 50. Only one generation of animals was simulated and no selection was considered.

Table 4.1. *Considered input parameters for simulation.*

Input parameter	Values
Number of animals per herd	10 & **50**[a]
Number of daughters per sire	25 & **100**
Use of sires across herds	**random**, selective and herd dependent
Residual variance (σ_e^2)	1 - σ_{level}^2
Correlation between level and slope	-0.5, **0** & 0.5
Variance for level (σ_{level}^2)	0.04 and **0.4**
Variance for slope (σ_{slope}^2)	0.02 and **0.2**

[a]Values in bold are default values.

Daughters of sires were either randomly or non-randomly assigned to herds following three different scenarios, based on differences in selection of sires and herds and resulting genetic connection between the groups of herds (Table 4.2). In the first scenario sires were assigned randomly across herds. In the second scenario (selective use of sires), sires were ranked based on simulated breeding value of level. Both sires and herds were split in five equally sized groups; sires based on ranking of their breeding values for level and herds at random. Daughters of sires from the first group were most likely assigned to herds of the first group; daughters of sires from the second group were most likely assigned to herds of the second group, etc. The chances of a sire from group *i* to have a daughter in group of herds *j*, are shown in Table 4.3. The third scenario involved non-random grouping of herds based on an increasing simulated herd effect combined with selective use of sires, to create a positive correlation between the herd effect and sires breeding values for level. This scenario is referred to as the herd dependent use of sires.

Table 4.2. *Different scenarios for use of sires, given the composition of groups of herds and sires and genetic connections between groups of herds.*

Use of sires	Groups of herds	Groups of sires	Genetic connection between groups of herds
Random	No groups	No groups	Strong
Selective	Random	TBV[a] of level	Poor
Herd dependent	Simulated herd effect	TBV of level	Poor

[a]True breeding value.

Table 4.3. *Chances that a daughter of a sire from one of the five groups of sires was assigned to a herd in one of the five groups of herds for selective use of sires.*

Group of sires	Group of herds				
	1	2	3	4	5
1	0.8318	0.1381	0.0247	0.0045	0.0009
2	0.1381	0.7080	0.1265	0.0229	0.0045
3	0.0247	0.1265	0.6976	0.1265	0.0247
4	0.0045	0.0229	0.1265	0.7080	0.1381
5	0.0009	0.0045	0.0247	0.1381	0.8318

Analysis of simulated data

The general model used to analyse the simulated data, with a linear random regression on a calculated EP, was:

$$y_{jk} = \mu + hr_j + \sum_{i=0}^{1} \alpha_{ik}\, p_{ij} + e_{jk}$$

where: y_{jk} is the performance of cow k; μ is the average for the trait across all animals; hr_j is either a fixed effect of herd j or a fixed polynomial regression common to all evaluated animals on phenotypic average within a herd (see below); $\sum_{i=0}^{1} \alpha_{ik}\, p_{ij}$ is the additive genetic effect of animal k in herd j where α_{ik} is coefficient i of the random regression on a polynomial (pol(x,t) option in ASREML) (Gilmour et al., 2002a) of environment of animal k and p_{ij} is element i of a polynomial resembling the calculated EP of herd j; e_{jk} is the residual effect of cow k in herd j.

Polynomials were used to rescale EP in order to facilitate convergence of the model. Estimated genetic variance matrix S had variances of level and slope on the diagonal and covariances between those on the off-diagonals. The estimated genetic variance in an environment with EP equal to EP1 was calculated as $\Phi_{EP1}S\Phi_{EP1}$', where Φ_{EP1} is a vector with polynomial coefficients of EP1 on each row. The estimated genetic covariance between environments with EP equal to EP1 and EP2, respectively, is calculated as $\Phi_{EP1}S\Phi_{EP2}$'. To compare results to simulated values, all estimates of genetic variance components were calculated back from the polynomial scale to the original scale per replicate and then averaged across replicates. ASREML (Gilmour et al., 2002a) was used for all analyses. For all situations considered, 50 replicates were simulated, which was sufficient to obtain reliable averages in initial test analyses.

Modelling of EP

Three models were considered for estimated herd effect (hr_j) and calculated EP:

Model 1. hr_j is a fixed effect of the herd as normally used in breeding value estimation models (Henderson, 1973) and EP was calculated as the average phenotypic performance of the trait within a herd.

Model 2. hr_j is a fifth order fixed polynomial regression common to all evaluated animals (Schaeffer and Dekkers, 1994) on EP, which was calculated as the average phenotypic performance of the trait within a herd.

Model 3. hr_j was a fixed effect of herd and EP was iteratively estimated with the general model. In the first iteration EP was equal to the average phenotypic performance of the trait in a herd. In all consecutive iterations EP was equal to the value of the fixed herd effect, estimated in the previous iteration. The iteration was stopped if all EP were equal to the values of the corresponding estimated fixed herd effects, *i.e.* the difference between each newly estimated fixed herd effect (hr_j) and EP

from the last iteration was smaller than the convergence criterion (a maximal absolute change of 0.001).

Model 3 was expected to remove possible bias from EP, resulting from non-random use of sires or low numbers of animals per herd. Model 3 resembled the simulation model most, since the calculated EP was equal to the estimated fixed herd effect. In situations where all three models were applied, a single data set was simulated in each replicate and analysed with each of the three described models.

Comparison of different methods to model EP

The effects of description of an EP were investigated by comparing estimated variance components, expected breeding values and calculated EP to simulated values for the different scenarios across all 50 replicates. Estimated variance components were used to calculate estimated genetic correlations of the trait expressed in different environments. Also, the correlations between TBV and EBV of sires were calculated for different values of EP to indicate problems arising from the selective use of sires when applying CF.

RESULTS
Variance components, breeding values and EP

Each replicate gave estimates of the residual variance, variances of level and slope and the covariance between level and slope. Averages and standard deviations of estimated variance components across the 50 replicates are shown in Table 4.4 for the low and the high heritability trait with r_{a_0,a_1} of 0.0 and random use of sires. The trends were generally the same for the low and high heritability trait. Variance components of models 1, 2 and 3 were hardly different. Estimated variances of the slope were underestimated for situations with 10 animals per herd.

Genetic correlations between level and slope for all situations considered in Table 4.4 were estimated on average 0.2 higher than simulated (results not shown). In replicates where the estimated correlation between level and slope became higher than 1, the (co)variance matrix was forced to be positive definite by fixing the correlation at 0.999 (Gilmour et al., 2002a). For the low heritability trait, the variance of the slope became very small in a considerable number of replicates leading to fixation of the correlation between level and slope at 0.999 and on average to a high estimate of the correlation between level and slope. For the high heritability trait, the overestimation of the correlation between level and slope mainly resulted from an overestimation of the covariance between level and slope.

In each replicate, values were calculated for EP for all herds and breeding values of level and slope were predicted for all animals. Average correlations between simulated herd effects and calculated EP, and simulated and expected breeding values of level and slope of sires, are given in Table 4.5 for the high heritability trait, $r_{a_0,a_1} = 0.0$ and random use of sires. Different

Table 4.4. *Estimated variance components for the different models, given different data structures, random use of sires, a low or high heritability trait and a simulated correlation between level and slope of 0.0.*

Trait h^2	Number of daughters per sire	Number of animals per herd	Model	$\sigma_e^{2\,(a)}$ (0.96 / 0.60) [b]	$\sigma_{level}^{2\,(a)}$ (0.04 / 0.40) [b]	$\sigma_{slope}^{2\,(a)}$ (0.02 / 0.20) [b]	$\sigma_{level,slope}^{(a)}$ (0.0) [b]	Covariance structures forced pd [c]
Low	25	50	1	$0.960_{0.009}$	$0.058_{0.026}$	$0.023_{0.021}$	$-0.010_{0.020}$	18
	25	50	2	$0.942_{0.009}$	$0.056_{0.025}$	$0.022_{0.021}$	$-0.008_{0.021}$	17
	25	50	3	$0.960_{0.009}$	$0.054_{0.027}$	$0.023_{0.020}$	$-0.008_{0.021}$	18
	100	50	1	$0.961_{0.009}$	$0.043_{0.018}$	$0.018_{0.010}$	$-0.001_{0.012}$	18
	100	50	2	$0.943_{0.009}$	$0.041_{0.017}$	$0.018_{0.010}$	$-0.001_{0.012}$	18
	100	50	3	$0.961_{0.009}$	$0.043_{0.018}$	$0.018_{0.010}$	$-0.001_{0.012}$	18
	100	10	1	$0.963_{0.009}$	$0.045_{0.012}$	$0.009_{0.008}$	$0.002_{0.008}$	16
	100	10	2	$0.873_{0.007}$	$0.035_{0.009}$	$0.006_{0.005}$	$0.004_{0.006}$	26
	100	10	3	$0.963_{0.009}$	$0.044_{0.012}$	$0.009_{0.008}$	$0.003_{0.008}$	18
High	25	50	1	$0.601_{0.019}$	$0.401_{0.037}$	$0.126_{0.036}$	$0.038_{0.033}$	0
	25	50	2	$0.600_{0.018}$	$0.383_{0.036}$	$0.124_{0.035}$	$0.037_{0.033}$	0
	25	50	3	$0.601_{0.019}$	$0.400_{0.038}$	$0.131_{0.036}$	$0.036_{0.034}$	0
	100	50	1	$0.608_{0.028}$	$0.403_{0.042}$	$0.134_{0.026}$	$0.029_{0.025}$	0
	100	50	2	$0.606_{0.026}$	$0.385_{0.041}$	$0.131_{0.025}$	$0.029_{0.025}$	0
	100	50	3	$0.608_{0.028}$	$0.402_{0.042}$	$0.140_{0.026}$	$0.026_{0.025}$	0
	100	10	1	$0.609_{0.025}$	$0.459_{0.032}$	$0.046_{0.013}$	$0.049_{0.013}$	1
	100	10	2	$0.593_{0.021}$	$0.365_{0.026}$	$0.037_{0.011}$	$0.049_{0.011}$	2
	100	10	3	$0.608_{0.025}$	$0.453_{0.032}$	$0.051_{0.015}$	$0.050_{0.015}$	1

[a] Standard deviations are given as a subscript. Standard error is equal to the standard deviation divided by $\sqrt{50}$.

[b] Simulated values for the low and high heritability trait, respectively.

[c] Positive definite.

definitions of EP hardly influenced correlations between simulated herd effects and calculated EP. The EP of models 1 and 2 were both calculated as phenotypic herd averages and therefore were the same. Generally, values of EP in model 3 converged after two or three iterations. The number of animals per herd had a larger effect on the correlations between simulated herd effects and calculated EP, than the number of daughters per sire. The number of daughters per sire had a larger effect on the correlations between simulated and expected breeding values of levels and slopes of sires, than the number of animals per herd.

Table 4.5. *Correlations between simulated herd effects and calculated environmental parameters (herd environment) and between simulated and estimated values of level and slope of breeding values of sires, given a high heritability trait, random use of sires, different data structures and a simulated correlation between level and slope of 0.0.*

Number of daughters per sire	Number of animals per herd	Model	Herd environment [a]	Level [a]	Slope [a]
25	50	1	$0.905_{0.005}$	$0.718_{0.009}$	$0.546_{0.016}$
25	50	2	[b]	$0.717_{0.009}$	$0.546_{0.016}$
25	50	3	$0.912_{0.005}$	$0.718_{0.009}$	$0.547_{0.016}$
100	50	1	$0.905_{0.006}$	$0.785_{0.015}$	$0.673_{0.028}$
100	50	2	[b]	$0.784_{0.015}$	$0.673_{0.028}$
100	50	3	$0.912_{0.006}$	$0.785_{0.015}$	$0.675_{0.028}$
100	10	1	$0.689_{0.008}$	$0.783_{0.020}$	$0.624_{0.030}$
100	10	2	[b]	$0.782_{0.020}$	$0.623_{0.031}$
100	10	3	$0.689_{0.008}$	$0.783_{0.020}$	$0.628_{0.029}$

[a] Standard deviations are given as a subscript.
[b] Environmental parameters used in models 1 and 2 are calculated in the same way, leading to the same correlation between simulated herd effects and calculated environmental parameters for models 1 and 2.

Genetic variances across environments estimated by model 1 are shown in Figure 4.2 for the high heritability trait with r_{a_0,a_1} equal to 0.0. Regardless of the data structure, the curve of the estimated genetic variance was flatter than the curve of the simulated genetic variance. The number of animals per herd had a strong influence on the estimates of the genetic variance, while the influence of the number of daughters per sire was limited. In the situation with 100 daughters per sire and 10 animals per herd, estimated genetic variance deviated up to 30% from the simulated value. The simulated value of $r_{g(EP=0.5,EP=1.5)}$ was 0.89 (given $r_{a_0,a_1} = 0.0$), while estimated values were 0.93, 0.93 and 0.97 (results not shown) for situations with 25 daughters per sire and 50 animals per herd, 100 daughters per sire and 50 animals per herd and 100 daughters per sire and 10 animals per herd, respectively.
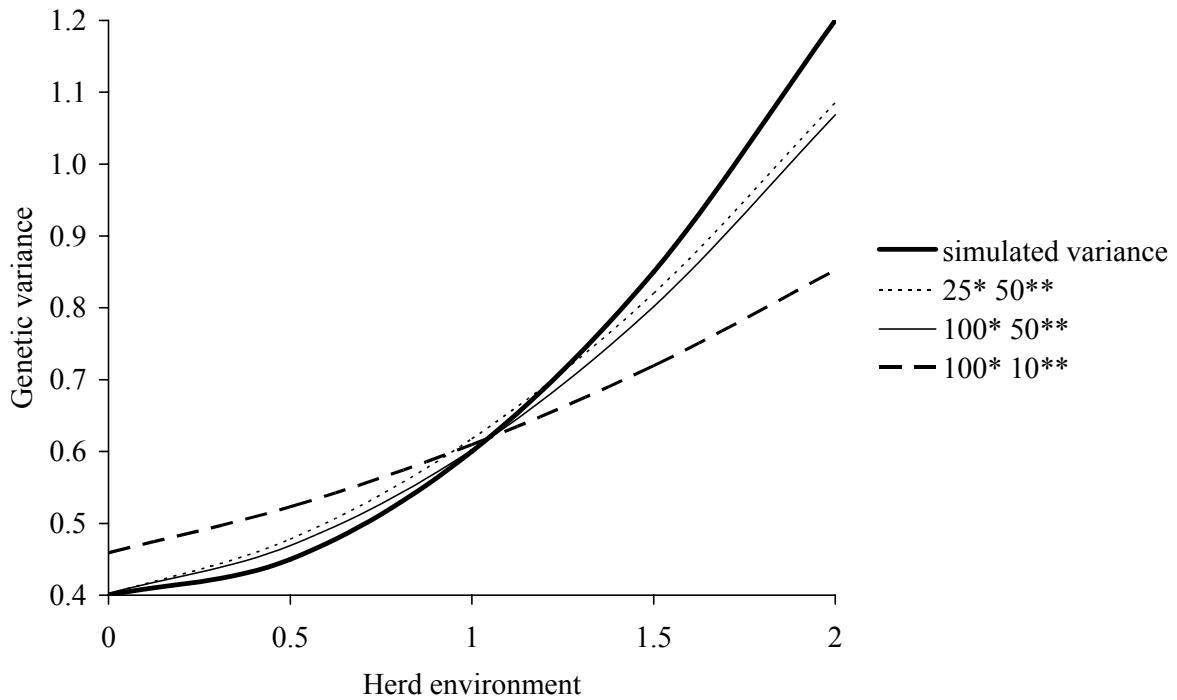
Figure 4.2. *Estimated (and simulated) genetic variance of the high heritability trait as function of the herd environment, given random use of sires and a correlation between level and slope of 0.0, for situations with 25 daughters per sire (*) and 50 animals per herd (**), 100 daughters per sire and 50 animals per herd and 100 daughters per sire and 10 animals per herd.*

Selective use of sires

Averages and standard deviations of estimated variance components of model 1 for selective use of sires are shown in Table 4.6. Residual variance was strongly overestimated and the variances of level and slope were strongly underestimated in all situations. For situations with selective use of sires, model 3 gave results (not shown) that were comparable to model 1, indicating that model 3 was not better in distinguishing between environmental and genetic effects than model 1.

Correlations between simulated herd effects and calculated EP and between simulated and expected breeding values of sires of level and slope are shown in Table 4.7. Correlations for herd environment and sires breeding values of level were lower than for situations with random use of sires, while correlations of the slopes of sires breeding values were slightly higher. Biased estimates of EP combined with underestimated variances of level and slope resulted in an underestimation of the genetic variance across environments in all situations with selective use of sires (results not shown).

Table 4.6. *Estimated variance components of model 1 for the high heritability trait, given selective use of sires and different data structures.*

Correlation level and slope	Number of daughters per sire	Number of animals per herd	$\sigma_e^{2\ (a)}$ $(0.60)^{(b)}$	$\sigma_{level}^{2\ (a)}$ $(0.40)^{(b)}$	$\sigma_{slope}^{2\ (a)}$ $(0.20)^{(b)}$	$\sigma_{level,slope}^{(a)(c)}$	Covariance structures forced positive definite
-0.5	25	50	$0.714_{0.010}$	$0.147_{0.029}$	$0.099_{0.025}$	$-0.034_{0.025}$	0
-0.5	100	50	$0.698_{0.015}$	$0.152_{0.027}$	$0.088_{0.022}$	$-0.020_{0.020}$	0
-0.5	100	10	$0.696_{0.014}$	$0.157_{0.020}$	$0.041_{0.009}$	$0.004_{0.010}$	0
0	25	50	$0.846_{0.016}$	$0.147_{0.024}$	$0.078_{0.021}$	$0.032_{0.018}$	2
0	100	50	$0.775_{0.028}$	$0.211_{0.038}$	$0.080_{0.018}$	$0.045_{0.016}$	2
0	100	10	$0.785_{0.029}$	$0.239_{0.038}$	$0.038_{0.012}$	$0.047_{0.012}$	3
0.5	25	50	$1.037_{0.017}$	$0.139_{0.027}$	$0.064_{0.021}$	$0.061_{0.020}$	20
0.5	100	50	$0.875_{0.040}$	$0.299_{0.064}$	$0.058_{0.015}$	$0.087_{0.017}$	11
0.5	100	10	$0.885_{0.042}$	$0.348_{0.055}$	$0.033_{0.009}$	$0.069_{0.010}$	9

[a] Standard deviations are given as a subscript.
[b] Simulated values.
[c] Simulated values of covariance between level and slope were –0.141, 0.0 and 0.141 for situations with correlations between level and slope of –0.5, 0.0 and 0.5, respectively.

Table 4.7. *Correlations between simulated herd effects and calculated environmental parameters (herd environment) and between simulated and estimated values of level and slope of breeding values of sires using model 1, given a high heritability trait, selective use of sires and different data structures.*

Correlation level and slope	Number of daughters per sire	Number of animals per herd	Herd environment [a]	Level[a]	Slope[a]
-0.5	25	50	$0.831_{0.010}$	$0.277_{0.016}$	$0.466_{0.023}$
-0.5	100	50	$0.827_{0.012}$	$0.500_{0.042}$	$0.511_{0.030}$
-0.5	100	10	$0.672_{0.009}$	$0.495_{0.028}$	$0.482_{0.030}$
0	25	50	$0.728_{0.013}$	$0.392_{0.014}$	$0.661_{0.015}$
0	100	50	$0.739_{0.018}$	$0.662_{0.033}$	$0.707_{0.031}$
0	100	10	$0.595_{0.014}$	$0.646_{0.037}$	$0.685_{0.028}$
0.5	25	50	$0.640_{0.020}$	$0.490_{0.017}$	$0.709_{0.013}$
0.5	100	50	$0.634_{0.022}$	$0.800_{0.019}$	$0.840_{0.018}$
0.5	100	10	$0.525_{0.013}$	$0.788_{0.026}$	$0.831_{0.018}$

[a] Standard deviations are given as a subscript.

Herd dependent use of sires, the situation with a confounding of sires breeding values of level and simulated herd effect, was only applied to the situation with 100 daughters per sire and 50 animals per herd with a correlation between level and slope of 0.0. The results (not shown) were comparable to the results for the selective use of sires.

Table 4.8. *Simulated and estimated average breeding values[a] for groups of sires 1, 3 and 5 in case of random, weak, strong or herd dependent selective use of sires given the high heritability trait, 100 daughters per sire and 50 animals per herd with a correlation between level and slope of 0.0.*

		Average breeding values			Correlations		
	Group	EP			EP		
Use of sires	of sires	0.57	1	1.43	0.57	1	1.43
Simulated[b]	1	$-0.886_{0.047}$	$-0.888_{0.058}$	$-0.890_{0.073}$			
	3	$-0.002_{0.042}$	$-0.003_{0.056}$	$-0.004_{0.072}$			
	5	$0.878_{0.051}$	$0.876_{0.065}$	$0.874_{0.082}$			
Random	1	$-0.756_{0.047}$	$-0.830_{0.049}$	$-0.905_{0.054}$	$0.890_{0.020}$	$0.928_{0.012}$	$0.920_{0.014}$
	3	$-0.002_{0.037}$	$-0.003_{0.043}$	$-0.003_{0.050}$	$0.885_{0.020}$	$0.910_{0.015}$	$0.911_{0.016}$
	5	$0.743_{0.048}$	$0.816_{0.052}$	$0.890_{0.059}$	$0.899_{0.020}$	$0.934_{0.013}$	$0.927_{0.015}$
Selective	1	$-0.497_{0.051}$	$-0.559_{0.051}$	$-0.621_{0.053}$	$0.902_{0.016}$	$0.935_{0.013}$	$0.927_{0.016}$
	3	$0.016_{0.033}$	$0.013_{0.038}$	$0.010_{0.043}$	$0.895_{0.020}$	$0.909_{0.018}$	$0.910_{0.018}$
	5	$0.463_{0.059}$	$0.531_{0.063}$	$0.599_{0.069}$	$0.898_{0.019}$	$0.931_{0.013}$	$0.924_{0.015}$
Herd	1	$-0.507_{0.063}$	$-0.560_{0.064}$	$-0.614_{0.067}$	$0.890_{0.023}$	$0.865_{0.025}$	$0.824_{0.030}$
dependent	3	$0.024_{0.046}$	$0.018_{0.052}$	$0.011_{0.059}$	$0.901_{0.017}$	$0.912_{0.014}$	$0.911_{0.015}$
	5	$0.442_{0.060}$	$0.510_{0.064}$	$0.578_{0.070}$	$0.862_{0.026}$	$0.939_{0.012}$	$0.951_{0.009}$

[a] Breeding values were calculated as the sum of level and EP×slope. Standard deviations are given as subscripts.
[b] Simulated values were averaged across the three situations.

Expected breeding values across environments

For the situation with 100 daughters per sire, 50 animals per herd, a correlation between level and slope of 0.0 and all three scenarios of selective use of sires, simulated and expected breeding values were calculated for three values of EP. Chosen values were median values of EP of groups of herds 1, 3 and 5 in the case of herd dependent use of sires. Averages and standard deviations of EBV across replicates are shown in Table 4.8. The group of sires 1 represented the 100 sires with the lowest simulated breeding values for level, the group of sires 3 represented the 100 sires with simulated breeding values for level around average and the group of sires 5 represented the 100 sires with the highest simulated breeding values for level. Averages of simulated breeding values of groups of sires in Table 4.8 were independent

from EP, due to the correlation between level and slope of 0.0. For groups of sires 1 and 5, $EBV_{EP=0.57}$ and $EBV_{EP=1.43}$ were on average closer to zero than simulated. As the data became more complex, average EBV of groups of sires 1 and 5 were closer to zero.

Correlations were calculated between simulated and expected breeding values for each EP level (Table 4.8). Correlations were slightly higher for EP = 1.00 and EP = 1.43. Correlations were the same for random and selective use of sires. For herd dependent use of sires, correlations tended to be the highest in the group of herds where sires had most daughters.

DISCUSSION

Modelling of EP

In this study we started with an idealised situation where the simulation model and the model used to analyse the data, were as similar as possible. One of the major differences between the simulation and estimation models was that EP in model 1 and 2 were calculated as phenotypic averages since they are generally modelled in a reaction norm model (Veerkamp and Goddard, 1998; Kolmodin et al., 2002; Calus and Veerkamp, 2003; Fikse et al., 2003a). The proposed alternative model (model 3) was expected to correct for genetic influences on EP by iteratively estimating the fixed herd effect in the evaluation model and use this as EP in the next iteration. Model 3 was tested because we expected that this model had closer resemblance with the simulated (and probably the true) model. All models used a linear random regression on EP to model genetic effects. Model 1 performed slightly better than model 2, which likely results from the fact that model 1 exactly fitted the simulation model and used more degrees of freedom to estimate herd effects. The results of model 3 were not different from results of model 1 even for the situation with selective use of sires and model 3 used about twice as much calculation time as model 1. Failure of the alternative model to perform better than model 1 could mean that either simply using herd means as EP is not the real underlying problem for estimation when using data under the scenario of selective use of sires, or that the proposed alternative model did not properly account for possible genetic bias in EP. More theoretical models, that for instance include simulated environmental effects as EP, could be used to explore the nature of this problem further. However based on the results of this study which was restricted to practical applicable models, there is no reason to use model 3 instead of model 1.

Model convergence was one of the major problems experienced with all three models. Although the random regression model is the most common applied covariance function in reaction norm models, Jaffrezic and Pletcher (2000) showed in a few examples that a character process model was more successful in modelling longitudinal data than random regression. The application of a character process model to model reaction norms appears straightforward, and might provide a solution to get better convergence of the model.

Estimation of G × E

For the trait with a low heritability it was more difficult to estimate the genetic CF than for the trait with a high heritability. The main problem was that for the low heritability trait in almost 40% of the replicates the covariance structure was forced to be positive definite. Also, the number of animals per herd was important to estimate genetic variance and calculate EP correctly, which illustrates that environmental sensitivity is better estimated in a population with larger herds and likely to be underestimated for a population with small herds. However, in a practical situation small herds may either be too large in number to simply disregard or represent certain management styles that are hardly found in larger herds. This problem might be partly solved by calculating EP based on for instance 50 animals that calved consecutively in one herd, rather than based on herd-year. Changing the data structure from the default situation by reducing the number of animals per herd to 10 or by introducing the non-random use of sires, led to correlations between simulated herd effects and calculated EP of 0.69 and 0.74, respectively (Tables 4.5 and 4.7). Although these changes in data structure are arbitrary, it indicates that both relatively low numbers of animals per herd and non random use of sires leads to biased EP.

One of the effects observed was that the estimated genetic correlation between the high heritability trait expressed in different environments was biased upwards, *i.e.* estimates were 0.93, 0.93 and 0.97 in situations with a random use of sires where the simulated value was 0.89. This resulted from the overestimated covariance between level and slope and the underestimation of variance of slope. Underestimation of variance of slope in situations with random use of sires also resulted in deviations of up to 30% of estimated genetic variance from simulated genetic variance. Variances of slope were more underestimated if the population structure was less informative, which indicates that high estimates of the genetic correlation between a trait expressed in different environments calculated with CF might result from the quality of the data rather than from the absence of re-ranking based on TBV. In the extreme situation where the variance of slope is estimated to be zero, the estimated genetic correlation between a trait expressed in different environments will be 1, since it can easily be derived from equation (1).

Prediction of breeding values across environments

One of the objectives was to investigate the influence of sires breeding values on EP. In situations with selective and herd dependent use of sires, sires were grouped based on their TBV for level, which is equal to $TBV_{herd=0}$. Grouping of sires based on TBV for any other simulated herd effect would have caused only small changes in the composition of groups of sires, since the simulated genetic correlation between the trait expressed in different environments was relatively high. The model clearly had more problems in estimating effects correctly in the case of selective use of sires. Selective use of sires not only implies a possible

bias in EP but also poorer genetic connections between groups of herds, which can lead to more difficulties for the model to disentangle genetic and environmental effects (Foulley et al., 1990). From this study, it is not clear whether problems in the estimation of variance components in the case of non random use of sires are due to genetic influence on EP, poorer genetic connections between groups of herds or failure to disentangle genetic and environmental effects. Random herd effects could be applied to avoid herd effects from absorbing part of the genetic levels within herds. Initial analyses with model 3 using random herd effects showed however that variances of level and slope were severely overestimated and herd variances were severely underestimated.

Groups of sires shown in Table 4.8 were selected based on their TBV of level. This implies that selection was based on data that is not included in the genetic evaluation and therefore EBV are expected to be biased (Henderson, 1973) and correlations between TBV and EBV are expected to be different for different groups of sires. However, groups of sires in Table 4.8 were the same for the different scenarios. Therefore, differences between scenarios are due to differences in genetic compositions of herds and in case of herd dependent use of sires also due to the fact that sires had most of their daughters in a limited range of environments. Correlations between simulated and expected breeding values indicated that breeding values of sires were predicted accurately across environments with the different models. Absolute values of EBV, however, were closer to zero if the data became less informative. This is not a problem if selection is based on a single trait or if scaling effects are not important. If selection is, however, based on an index based on more than one trait with different scaling effects, scaling effects can cause re-ranking across environments based on the composite index (Namkoong, 1985). In that case, non-random use of sires could result in misleading indexes, since scaling effects of traits are likely to be underestimated.

The EBV of cows were not compared to their TBV. In the simulated data, cows only had one record and therefore only one point through which their EBV was fitted. Since the EBV of cows are based on far less data than the EBV of sires, the EBV of cows are likely to be more biased than the EBV of sires, especially if breeding values are extrapolated to extreme environments.

Problems in estimating variance components and lower correlations between TBV and EBV under the presence of selective use of sires seem to contradict suggestions (Kolmodin et al., 2002) that CF could be useful in overcoming problems with genetic connectedness in international breeding value estimation. However, only one population with one generation of sires was simulated and subsets of sires were not equally distributed, while an international situation ideally would be simulated by different related base populations reflecting different countries. Additional genetic relations between animals would improve genetic connectedness across environments and therefore reduce bias in EBV. Since poor genetic connectedness and confounding between herd and genetic effects are features of the data, bias in estimated

variance components might be reduced by selecting data containing genetically well-connected herds with a non-extreme genetic composition and different levels of management.

CONCLUSION

Implications of using phenotypic averages as EP in CF were expected to lead to problems of estimation of variance components. Non average genetic composition of herds and poor genetic connectedness had a large impact on estimated variance components in CF and gave poorer correlations between simulated and predicted sire effects and between simulated herd effects and calculated EP. Estimation problems were not overcome by a new model that aimed at separating environmental and genetic effects in the EP. The effect of estimation problems was that genetic correlations between the trait expressed in different environments were biased upwards and that EBV were biased if genetic connectedness became poorer and herd composition more diverse. The best possible solution at this stage is to use EP combining a large number of animals per herd.

ACKNOWLEDGEMENTS

Chapter 5

Associations among descriptors of herd management and phenotypic and genetic levels of health and fertility

M. P. L. Calus

J. J. Windig

R. F. Veerkamp

Animal Sciences Group
Division Animal Resources Development
P.O. Box 65, 8200 AB Lelystad, The Netherlands

ABSTRACT

The objective of this paper was to investigate the association of descriptors of herd environment with phenotypic levels and breeding values of fertility and health traits. Analyses were performed for 82,080 first-lactation heifers and 173,787 multiparous cows. Fourteen environmental parameters were defined, that described herd environment, such as average protein production, average somatic cell score (SCS), average calving interval, and average body condition score (BCS). Herds with lower average SCS had, in general, more desirable values for almost all analyzed traits (i.e., days to first service was 7 d shorter), as did herds with lower average calving interval (i.e., 2.8% lower incidence of predicted mastitis). Herds with higher average protein production had slightly poorer fertility but more desirable values for all other analyzed traits (i.e., 5.1% less predicted mastitis, 0.4 lower SCS and 0.6 higher BCS). Variance components and breeding values of sires were estimated by applying a random regression on the environmental parameters. In general, genetic variances varied only slightly across environments. However, based on data exclusively from heifers, the genetic variance for number of inseminations was 4.1 times higher in herds with a higher number of inseminations, 1.9 times higher for survival in herds with higher fat-to-protein ratio, and 1.7 times higher for predicted mastitis in herds with higher number of inseminations. Based on the heifer data, the lowest estimated genetic correlation across environments was 0.76 (SE 0.21) for first-service conception between herds with differing average BCS. The minimum based on the cow data was 0.65 (SE 0.10) for survival between herds with differing average ages at calving. The relative importance of some fertility traits compared with yield traits doubled across environments. Possible reranking of individual animals within a population and the changes in genetic variance across environments suggests that environment-specific breeding values should be estimated for use in customized selection indices.

INTRODUCTION

Possible options to compensate for the increasing fertility and health risks associated with selection for increased milk production (Emanuelson, 1988; Pryce et al., 1998; Rauw et al., 1998) are: 1) improvement of management such that the poorer genetic merit is alleviated, or 2) inclusion of health and fertility in genetic selection with sufficient weight to ensure there is no reduction of the genetic level for health and fertility (Philipsson et al., 1994). The extent to which these 2 options interact is often labeled as genotype-by-environment interaction, environmental sensitivity of genetic variance (**ES**) or genetic variance in reaction norms. If the effects of genotype-by-environment interactions are important relative to the average effects of genotype and environment, ignoring management improvement might hinder expression of the genetic improvement for health in poor environments. Alternatively, it could mean that observed genetic differences between animals are larger in poor health environments, and hence, there might be more benefit of genetic selection in these

environments. Therefore, in the presence of ES, management and genetic improvement of health and fertility might support or counteract each other. Other consequences for genetic selection are that the relative importance of traits might change across environments, and thus the weights in total merit indexes (Namkoong, 1985), and that available EBV (theoretically most applicable to an average environment) might not be sufficient to select animals for specific herd environments. The consequences of ES for improvements in health and fertility due to reasons other than genetic selection might be that the expected response depends on the genetic background of the animals.

An interesting implication of ES is that selection on high phenotypic performance combined with a continually improving herd environment is expected to increase ES of the animals, as indicated in a simulation study (Kolmodin et al., 2003). Thus, in the long-term, improving management to alleviate the lower genetic level for health and fertility might result in a continuously smaller range of environments where animals maintain their health and fertility. This means that the expected increase in ES of the animals increases the importance of tuning genotype and environment. In the long-term, including existing ES in multitrait genetic selection might be necessary to enable specific selection of animals for a wide range of herd environments.

Studies about ES of health and fertility traits used different methods, such as regression of health traits on pedigree indexes for production (Pryce et al., 1999), including a sire-by-herd interaction term in the statistical model for SCS (Samore et al., 2001), or estimating a genetic correlation for SCC between environments with a low or high within-herd standard deviation for milk yield (Castillo-Juarez et al., 2000; Raffrenato et al., 2003) or low or high herd-year average SCS (Banos and Shook, 1990). Those studies did not report significant ES. Reaction norm models including a genetic covariance function describing (co)variances over the range of environments (Kirkpatrick and Heckman, 1989) were applied for fertility traits in only a few studies, where heterogeneous heritabilities across fertility environments were reported (Distl, 2001; Kolmodin et al., 2002). Covariance functions enable us to include both heterogeneous genetic variances and genetic correlations as function of a continuous measure of the environment, avoiding arbitrarily grouping of environments, and might therefore provide a better method to model ES.

The objective of this paper was to investigate the association of several descriptors of herd environment with 1) phenotypic levels of fertility and health, and 2) ES of breeding values for fertility and health obtained using a reaction norm model.

MATERIALS AND METHODS
Test day and insemination data
Records for Insemination and yields of milk, fat, and protein yield and SCC were available for 147,835 first lactation heifers and 295,507 multiparous cows calving between July 1997

and June 1999. All animals were at least 75% Holstein-Friesian. Years were defined from July 1 through June 30, to ensure that the months with most calvings fell in the middle of the defined year. First-lactation heifers were selected if they calved on an age between 640 and 1095 d. Editing steps for environmental parameters (see below) reduced the number of heifers and cows to 116,727 and 230,887, respectively. For heifers and cows separately, all (grand)daughters of (grand)sires with less than 20 (grand)daughters in the edited data were deleted, reducing the number of heifers to 87,375 and the number of cows to 192,615. Herd-year-season subgroups were formed based on the method of Crump et al. (1997) with a minimum of 5 animals per subclass, a minimum length of 30 d, and a maximum length of 365 d. Records of animals that could not be assigned to a group with at least 5 records or that were assigned to a group with fewer than 3 informative records for any of the traits were deleted. Additionally, (maternal grand)sires with progeny in fewer than 3 herd-year-season classes, and herd-year-season classes with progeny of less than 3 (maternal grand)sires were deleted. The editing steps based on herd-year-season subclasses and offspring per sire were repeated until the final data set met all criteria, reducing the number of records of heifers to 82,080 and the number of records of cows to 173,787. For the different traits, between 61,002 (calving interval) and 79,068 (SCS and predicted mastitis) records were informative for the heifer data, and between 118,818 (calving interval) and 167,031 (SCS and predicted mastitis) records were informative for the cow data.

BCS data

Body condition scores were available for 76,811 heifers, of which 12,823 calved between July 7, 1997 and June 30, 1998, and 63,988 calved between July 1, 1998 and June 30, 1999. The BCS was scored by classifiers during herd classification. Records of animals scored after 305 DIM were deleted, reducing the number to 74,554. Herd-year-season subclasses for BCS were defined as herd-visits of the classifiers. Animals in herd-year-season subclasses with fewer than 5 animals were deleted. These editing steps, combined with the criteria for environmental parameters (see below) reduced the number of records for BCS to 68,418 in 6184 herd-years. Heifers that calved in those herd-years and had a record for at least one of the other traits but no BCS records were included in the data with a missing value for BCS. This increased the number of records to 85,631. Herd-year-season subgroups for all other traits were formed based on the method of Crump et al. (1997) with a minimum of 5 animals per subclass, a minimum length of 30 d and a maximum length of 365 d. (Grand)daughters of (grand)sires with less than 10 (grand)daughters in the data were deleted. In total, 69,906 records were included for the analyses where BCS was included as trait or environmental parameter. For the different traits, between 50,653 (calving interval) and 66,923 (SCS and predicted mastitis) records were informative.

Pedigree

Initially all sires, paternal grand dams and maternal grandsires of animals with records in the data were included in the pedigree file. All male predecessors of those animals, available from the pedigree data, were included. Identification of dams of bulls was included if a dam had 2 or more sons; otherwise, dams were included as base parents. For the heifers, in total 1754 (2361 for BCS) animals were included in the relationship matrix. For the cows, 3442 animals were included in the relationship matrix.

Traits

Fertility traits. Seven fertility traits were considered: days to first service (**DFS**), days to last service (**DLS**), days first to last service (**DFLS**), calving interval (**CIV**), number of inseminations per service period (**NINS**), first service conception (**FSC**), and non-return at 56 d after first insemination (**NR56**). The DFS was calculated as interval from calving to first service, DLS as interval from calving to last service, and DFLS as interval from first to last service. The CIV was the interval between 2 consecutive calvings. The NINS was the number of inseminations per service period. The FSC was 1 if the cow had only one insemination and a known next calving date, and 0 otherwise. The NR56 was 1 if within 56 d after the first insemination no second insemination was recorded and 0 otherwise.

Records for any of the traits were missing if no information was available to calculate the value for the trait. Records for DFS were missing if DFS was smaller than 20 or greater than 300. Records for DLS were missing if DLS was smaller than 20 or greater than 500. Records for DFLS were missing if DFLS was greater than 400. The NINS was missing if NINS was 0 or greater than 10. Records for CIV were missing for animals without a known next calving date, or if CIV was smaller than 300 or greater than 800. These criteria were applied to exclude extremely long lactation records, records with extreme short gestations due to abortions, or records with errors.

Survival. Survival was defined following Pool et al. (2003), being 1 for cows with known next calving date. Survival was coded as 0 for cows without a known next calving date and with the last test-day record occurring at least 140 d before the last recorded test day for the respective herd, as it was unlikely that a cow was still on the farm when no test-day records have been recorded in a period of 140 d. Survival of animals was missing in all other situations.

SCS and predicted mastitis. Somatic cell score was defined as the average SCS across test days. A binary trait, called predicted mastitis, was used as indicator trait for mastitis following De Haas et al (2004), being 1 if SCC on at least one test day during the lactation was greater than 400,000 cells/mL and 0 otherwise.

BCS. Body condition score was measured on a scale from 1 to 9 [thin to fat; based on Lowman et al. (1976)]. An average BCS curve across test days was fitted with a smoothing

spline (Gilmour et al., 2002b) based on all available records. The deviance from the average BCS curve across DIM was used for the analyses.

Environmental parameters

Fourteen different environmental parameters (**EP**), describing herd management, were calculated as an average from all animals that had information on the characteristic and calved in the same herd-year. The EP were herd-year averages of protein production, fat-to-protein ratio, SCS, persistency, relative peak milk yield, age at calving, number of animals, change in the number of animals between consecutive herd-years, change in fat percentage between 14 and 77 DIM, calving interval, number of inseminations, peak calving date, distribution of calving dates, and BCS. For each individual EP, all available information was included and at least 25 animals in a herd-year needed to be informative for the characteristic. For average BCS and CIV, this criterion was, respectively, 5 and 10 animals, to prevent loss of great numbers of animals. For the same reason, no restriction was put on the EP average CIV in the BCS data. The EP were chosen because they represented management and herd environment, being, for instance, indicators for herd-year levels of production, energy balance, and fertility. More detailed reasoning behind the selection of applied EP, grouping of EP, and full description of the calculation of these EP is given by Calus and Veerkamp (2003).

Estimation of mean phenotypic performance across environments

To estimate the relation between the mean phenotypic performance of the animals for the considered traits and the values of the EP, a model was used that corrected for possible systematic effects influencing the mean phenotypic performance. The model included fixed linear and quadratic regressions for age at calving and breed, and the relationship between the mean phenotypic performance and the EP was modeled with a 10th-order polynomial regression on EP. The same fixed effects were included in the model to estimate ES (for details see below). The relative change in mean phenotypic performance (Δmpp) across environments was calculated as Δmpp = {(mpp90th − mpp10th) / mpp50th} × 100%, where mpp10th, mpp50th, and mpp90th are mean phenotypic performances at 10th, 50th, and 90th percentiles of the data ordered on increasing values of the analyzed EP.

Estimation of variance components and environmental sensitivity

Variance components were estimated separately for first-lactation heifers and multiparous cows with a sire-maternal grandsire model. Fixed effects were included in the model for mean, parity (only for the multiparous cows), and herd-year-season subclass. Fixed regressions were included to account for age at calving and for breed of the cow. A 10th-order fixed polynomial regression on EP was included, to account for the average effect across EP. The ES was modeled by applying a random regression for each (maternal grand)sire,

representing its EBV, on values of an EP for the herd-years in which its (grand)daughters were producing. The incidence matrix of maternal grandsire effects was laid over the matrix of sire effects, i.e., if a bull had both entries in the data as sire and maternal grandsire, the breeding value as maternal grandsire was equal to half the breeding value as a sire. A random permanent environmental effect was included for the multiparous cows. The residual variance was estimated separately for 5 equally sized groups, based on increasing EP, to include heterogeneous residual variances in the model.

The applied model was:

$$Y_{klmno} = \mu + FIXED\ EFFECTS + \sum_{i=0}^{10} \beta_i P_{ik} + \sum_{j=0}^{s} \alpha_{jl} P_{jk} + 1/2 * \sum_{j=0}^{s} \alpha_{jm} P_{jk} + pe_o + E_{klmno}$$

where Y_{klmno} is the performance of cow o; μ is the average performance over all animals; *FIXED EFFECTS* included herd-year-season subclasses, parity (only for the multiparous cows: 2,3,4+), and second order polynomial regressions on age at calving and percentage of Holstein Friesian, Dutch Friesian and Meuse-Rhine-Yssel genes; β_i is coefficient i of a fixed regression on element i of the polynomials of all environments; P_{ik} is element i of the 10th-order polynomial of an environmental parameter of environment k; α_{jl} is coefficient j of the random regression on the orthogonal polynomials of all environmental parameters of the daughters of sire l; P_{jk} is element j of the orthogonal polynomial resembling an environmental parameter of environment k; α_{jm} is coefficient j of the random regression on the orthogonal polynomials of all environments of the maternal granddaughters of sire m; s is the largest significant estimable coefficient j of the random regression for sire effects; pe_o is a permanent environmental effect of cow o (only for the multiparous cows for all traits except survival); and E_{klmno} is the residual effect of cow o in environment k within group of environments n (n = 1, 2, ..., 5).

Definition of the genetic model resulted in estimated sire variances as a function of the values of the EP. Heritabilities were calculated as 4 times the sire variance divided by the sum of the residual variance (and the permanent environmental variance for the multiparous cows) and 1.25 times the sire variance. The factor 1.25 is explained by the fact that both effects for sires (1 times the sire variance) and maternal grand sires (0.25 times the sire variance) explain part of the genetic variance. All analyses were performed with ASREML (Gilmour et al., 2002b). Residual covariances between groups of environments were assumed to be zero. All combinations of EP and traits were tested for appearance of ES, using the likelihood ratio test to identify the highest estimable significant order for the sire effect ($P < 0.05$). The test statistic was twice the difference in log likelihood between models with order n and n-1, respectively.

RESULTS

Environmental parameters

The mean, standard deviation, and the range for the EP used in the analysis of the heifer data are given in Table 5.1. The values for the EP used for analysis of the higher parity cows were similar. Environmental parameters had correlations between –0.40 and 0.27 among each other, except for a correlation of 0.84 between average persistency and relative peak milk yield. This result indicated that herds with a high persistency (as defined here) also tended to have a higher relative peak milk yield. Pairs of EP with highest correlations were calculated from the same traits.

Mean phenotypic performance across environments

The relationship between the value of the EP and the mean phenotypic performance of traits is given by the estimated fixed polynomial regression on EP. The results are applicable for the "average" heifer in the data, being 0.9% Meuse-Rhine-Yssel, 4.7% Dutch Friesian and 93.6% Holstein-Friesian, calving at an age of 791 d. The relative change of the trait means from the 10th to the 90th percentile of the data (Table 5.2), for DFS, DLS, and DFLS were 8.3, 10.4, and 13.5% (7, 13, and 5 d) respectively, with increasing average SCS. The NINS

Table 5.1. *Mean, standard deviation, range, and values at 10th and 90th percentiles of the environmental parameters for the heifer data.*

Environmental parameter	Mean	SD	10th	90th	Minimum	Maximum
Protein (kg/305 d)	289	27.0	256	323	167	408
Fat / protein	1.27	0.04	1.21	1.32	1.12	1.48
Somatic cell score	2.63	0.48	2.02	3.22	0.94	6.06
Persistency	1.54	0.12	1.40	1.69	1.15	2.35
Relative peak milk yield	1.40	0.04	1.35	1.46	1.27	1.73
Age at calving (d)	1370	134	1229	1580	947	1989
Number of animals	53.3	22.4	32	80	25	246
Change number of animals	3.2	9.0	-7	13	-49	57
Change in fat percentage (%)[1]	-0.43	0.20	-0.68	-0.17	-1.35	0.19
Calving interval (d)	392	16.9	373	412	345	527
Number of inseminations	2.07	0.35	1.68	2.52	1.18	4.28
Peak calving date (d)[2]	149	56	81	205	-10	414
Distribution of calving dates	0.35	0.10	0.24	0.48	0.00	0.93
Body condition score (scale 1-9)	4.35	0.70	3.46	5.22	1.46	7.44

[1]The average difference in fat percentage on the test-days closest to 77 and 14 DIM.

[2]Years were defined from July 1st until of June 30th. -10 means day 355 of the year before, 414 means day 49 of the next year.

increased 4.7 % (0.09 inseminations) with increasing average protein production and number of animals, and decreased 4.1% (0.09 inseminations) with increasing peak calving date. The FSC and NR56 decreased respectively 4.0 and 6.9%, respectively, (absolute decreases of 1.9 and 4.4%) with increasing average protein production and number of animals, and increased respectively 6.8 and 4.4% (absolute increases of 2.9 and 2.5%) with increasing peak calving date. Survival and CIV did not have distinct associations with any of the EP. Incidence of predicted mastitis decreased 18.5% (-5.1% incidence) with increasing average protein production and 20.9% (-6.1% incidence) with increasing change in fat percentage and increased 13.6% (+3.2% incidence) with increasing relative peak milk yield and 11.8% (+2.8% incidence) with increasing calving interval. The SCS decreased 14.1% (0.4) and BCS increased 13.8% (0.6) with increasing average protein production.

Table 5.2. *The relative change in mean phenotypic performance (Δmpp in %) from the 10th to the 90th percentile of the data, compared to the 50th percentile, as predicted by the 10th order regression model. In the last row the means across all environments are given for each trait.*

Environmental Parameter	Trait[1]										
	DFS	DLS	DFLS	CIV	NINS	FSC	NR56	SUV	PM	SCS	BCS
Protein	-2.8	-2.7	-1.9	-1.0	4.7	-4.0	-6.9	0.8	-18.5	-14.1	13.8
Fat / protein	2.6	0.0	-4.3	0.1	-2.1	4.7	2.1	0.5	-3.9	1.0	-2.8
Somatic cell score	8.3	10.4	13.5	3.1	0.0	-2.4	3.2	-2.6	80.4	54.5	-3.3
Persistency	-9.9	-8.2	-4.4	-2.7	2.4	-4.4	-5.1	-0.1	7.0	6.9	1.6
Rel. peak milk yield	-8.6	-7.6	-4.5	-2.6	1.0	-1.0	-2.6	0.4	13.6	9.2	-1.4
Age at calving	3.2	3.3	4.9	1.2	1.3	2.1	1.0	2.7	-7.2	-2.9	-4.1
Nr. of animals	-3.9	-0.8	6.5	-0.3	4.7	-6.1	-3.9	1.1	-4.7	0.4	3.2
Change nr. animals	1.5	3.6	7.9	1.0	1.6	-4.9	-1.1	2.9	-2.1	-0.7	-0.1
Change in fat %	-4.4	-4.1	-3.0	-1.3	-0.8	-0.9	-0.5	2.4	-20.9	-5.0	-5.2
Calving interval	22.4	25.0	29.5	7.8	6.9	-7.5	1.8	-0.9	11.8	5.2	
Nr. inseminations	-11.9	9.8	78.2	2.0	48.2	-37.2	-28.6	1.0	2.3	-2.3	3.2
Peak calving date	4.4	1.1	-4.9	0.4	-4.1	6.8	4.4	-0.1	9.2	5.1	-3.8
Distr. calving dates	-3.3	-2.4	-0.5	-0.8	0.7	-2.9	-3.0	-0.6	-3.8	-0.9	1.2
Body cond. score	-6.6	-6.1	-3.2	-1.9	0.6	-3.3	-4.2	0.5	-9.4	-2.2	40.5
Trait mean	87.3	128.4	40.7	403.5	2.07	0.45	0.59	0.80	0.25	2.23	4.36

[1]DFS = days to first service, DLS = days to last service, DFLS = days first to last service, CIV = calving interval, NINS = number of inseminations before conception, FSC = first service conception, NR56 = non-return at 56 days, SUV = survival, PM = predicted mastitis based on test-day SCC, SCS = lactation average SCS.

Environmental sensitivity

Significant ES was estimated in respectively 13.6 and 15.4% of all combinations of traits and EP for the heifers and the multiparous cows, respectively. Nearly all estimated ES was based on linear random regressions, but significant quadratic random regressions were fitted for the heifer data for survival combined with fat-to-protein ratio and change in fat percentage, and for the cow data for SCS combined with protein production. Of all combinations of EP and traits with significant ES based on the heifer data, estimated genetic correlations between the trait expressed in the 10th and 90th percentiles of the data ranged from 0.76 to 1.00, but most were close to unity (Table 5.3). The lowest estimates were 0.76 (SE 0.21) for FSC combined with average BCS, 0.83 (SE 0.10) for survival combined with change of number of animals, and 0.84 (SE 0.12) for survival combined with change in fat percentage. Of all combinations of EP and traits with significant ES based on the cow data, estimated genetic correlations between the trait expressed in the 10th and 90th percentiles of the data ranged from 0.65 to 1.00, but most were close to unity (Table 5.4). The lowest estimates were 0.65 (SE 0.10) for survival combined with average age at calving, 0.92 (SE 0.06) for DFLS combined with average SCS and 0.92 (SE 0.05) for CIV combined with average protein.

The estimated genetic variances at the 10th and 90th percentiles of the heifer data, as a ratio of the estimated genetic variance at the 50th percentile of the data, are given in Table 5.5 for all combinations of traits and EP with significant ES based on the heifer data. The genetic variance in the 90th percentile compared to the 10th percentile (of the heifer data) was, for instance, 4.1 times higher for NINS (EP = number of inseminations), 1.9 times higher for survival (EP = fat-to-protein ratio) and 1.7 times higher for predicted mastitis (EP number of inseminations). This means that genetic variance of NINS was greater in herds with more inseminations, genetic variance of DFLS was greater in herds with more inseminations, genetic variance of survival is larger in herds with higher fat-to-protein ratio and genetic variance of DLS was greater in herds with increased CIV. The genetic variances for the cow data increased up to 2 times between the 10th and 90th percentiles (results not shown).

In most cases, heritabilities changed little across environments. For the heifer data, heritabilities for DFS, DLS, DFLS, CIV, and NINS in the average environment were 0.09, 0.06, 0.03, 0.05, and 0.03, respectively. Heritabilities for the binary traits FSC, NR56, survival, and predicted mastitis were in the average environment respectively 0.01, 0.01, 0.03, and 0.07. The SCS and BCS had heritabilities in the average environment of 0.19 and 0.40. The largest relative change in heritability was for the trait survival combined with the EP fat-to-protein ratio, being 0.025 in the 10th and 0.048 in the 90th percentile of the data.

Table 5.3. *Estimated genetic correlations[1] between a trait expressed in the 10th and 90th percentiles of the heifer data for all combinations of environmental parameters and traits with significant environmental sensitivity.*

Environmental Parameter	Trait[2]										
	DFS	DLS	DFLS	CIV	NINS	FSC	NR56	SUV	PM	SCS	BCS
Fat / protein								0.89	0.95		
Somatic cell score										0.97	
Rel. peak milk yield										1.00	0.99
Age at calving											0.99
Nr. of animals											0.97
Change nr. of animals							0.83				0.96
Change in fat %	0.97		0.99				0.84				0.98
Calving interval	0.99	0.97									
Nr. inseminations			0.99		0.96				0.98		
Peak calving date									0.96		
Distr. of calving dates											1.00
Body condition score	0.94					0.76					1.00

[1]Standard errors of the genetic correlations ranged from 0.00 to 0.21.
[2]DFS = days to first service, DLS = days to last service, DFLS = days first to last service, CIV = calving interval, NINS = number of inseminations before conception, FSC = first service conception, NR56 = non-return at 56 days, SUV = survival, PM = predicted mastitis based on test-day SCC, SCS = lactation average SCS.

Table 5.4. *The genetic correlations[1] estimated on the cow data between 10th and 90th percentiles of the data for all combinations of environmental parameters and traits with significant genotype by environment interaction.*

Environmental parameter	Trait[2]									
	DFS	DLS	DFLS	CIV	NINS	FSC	NR56	SUV	PM	SCS
Protein				0.92				0.97	0.97	0.98
Somatic cell score			0.92						0.95	0.93
Age at calving					0.93			0.65		
Nr. of animals								0.98		
Calving interval	0.99	0.94		0.93	0.99					
Nr. inseminations		0.95	0.95	0.94	0.94					1.00
Distr. of calving dates						0.95				

[1]Standard errors of the genetic correlations ranged from 0.01 to 0.10.
[2]See table 5.3.

Table 5.5. *The relative genetic variance estimated at 10th and 90th percentiles of the heifer data (the genetic variance at the 50th percentile is set to 1.0) for all combinations of environmental parameters and traits with significant genotype by environment*

Environmental parameter	DFS		DLS		DFLS		NINS		FSC		SUV		PM		SCS		BCS	
	10th	90th	10th	90th	10th	90th	10th	90th	10th	90th	10th	90th	10th	90th	10th	90th	10th	90th
Fat / protein											0.72	1.38	1.11	0.96				
SCS															0.89	1.15		
Peak milk yield															1.01	0.99	1.08	0.84
Age at calving																	1.29	1.01
Nr. of animals																	1.14	1.05
Change nr. anim.											1.08	1.10					1.09	0.95
Change in fat %	0.96	1.07			1.29	0.79					0.90	0.84					1.05	0.89
Calving interval	0.78	1.38	0.77	1.41														
NINS					0.60	1.68	0.49	1.99					0.80	1.32				
Peak calving date													1.18	0.89				
Distr. calv. dates																	0.90	0.99
BCS	1.15	0.94							0.98	1.41							0.72	1.26

[1]DFS = days to first service, DLS = days to last service, DFLS = days first to last service, CIV = calving interval, NINS = number of inseminations before conception, FSC = first service conception, NR56 = non-return at 56 days, SUV = survival, PM = predicted mastitis based on test-day SCC, SCS = lactation average SCS.

Breeding values for survival for the 10 sires with most daughters in the heifer data, estimated as function of herd-year average fat-to-protein ratio, followed different patterns across environments (Figure 5.1). The difference in survival of the first lactation of daughters of 2 particular sires (indicated with squares or triangles) was 2.9% in herd environments with a fat-to-protein ratio of 1.19 and -1.0% in herd environments with a fat-to-protein ratio of 1.33.
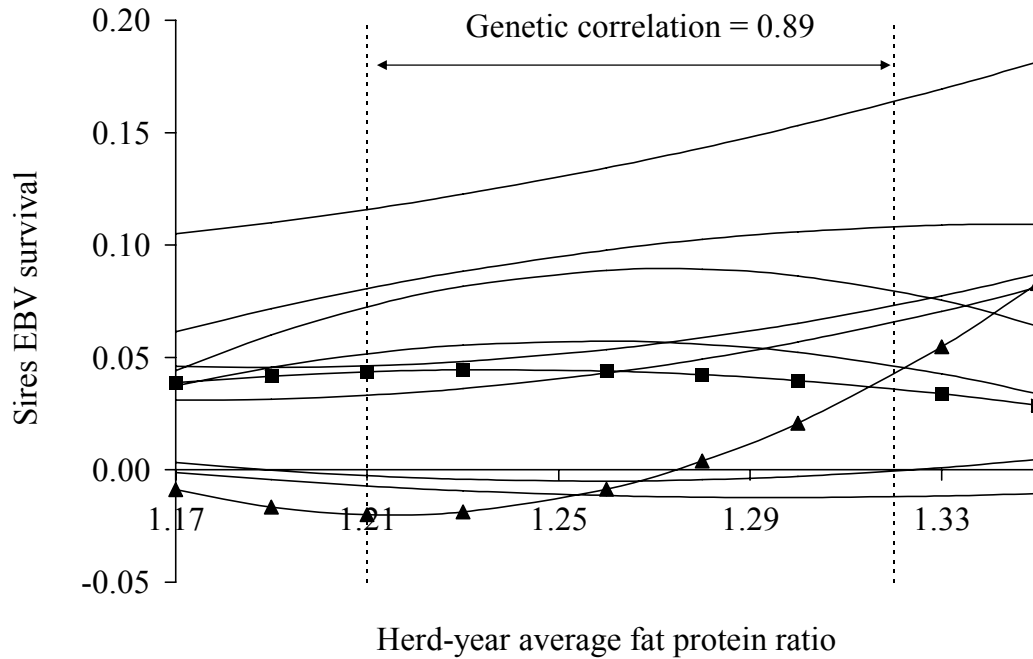

Figure 5.1. *Breeding values for survival of the 10 sires with most daughters in the heifer data, estimated as function of herd-year average fat-to-protein ratio (squares and triangles mark breeding values of 2 particular sires). Tenth and 90th percentiles of the data are shown as dotted lines.*

DISCUSSION

The objective of this paper was to investigate associations among herd environment and fertility and health at both the phenotypic and genetic level. Herd average SCS and herd average CIV generally had a stronger association with average cow performance than did average protein production, indicating that other descriptors of herd management might be more important than production level per se. Although significant ES was only detected in a limited number of situations, the most extreme (i.e., the lowest) genetic correlation of a trait expressed in different environments was 0.7, and genetic variance of some traits increased up to 4 times across environments. It is also important to note that the changes in genetic variance were in some situations in the opposite direction with respect to changes in genetic variance of production traits.

Herd characteristics explaining different levels of health and fertility

For the traits DFS, DLS, DFLS, CIV, NINS, predicted mastitis, and SCS, low values are desired, whereas high values are desired for the traits FSC, NR56, survival, and BCS. Lower herd averages for SCS and CIV were associated with more desirable average phenotypic values of almost all analyzed traits. A likely explanation might be that herds with lower SCS and shorter average calving interval have on average better management, resulting in better overall fertility and health. This hypothesis is in agreement with results of Rougoor et al. (1999), who concluded that farmers that are aware of their bulk milk SCC and average CIV have better hygiene, lower feed costs, and higher milk price, resulting in a higher gross margin. Moreover, the results of Barkema et al. (1999b) indicated that lower bulk milk SCC was associated with farmers that worked more precisely, paid more attention to individual cows, and made a greater effort to prevent mastitis.

Herds with higher average protein production had slightly lower success of insemination, but considerable less mastitis, lower SCS, and higher BCS. This result indicates that higher levels of production do not necessarily lead to poorer performance in other traits, possibly due to superior general management in those herds. The observation that heifers had less mastitis and lower SCS on farms with higher overall protein production, seems to contradict reported positive within-herd phenotypic correlations between milk yield and mastitis (Ingvartsen et al., 2003). An explanation might be that herds with high production levels manage to control SCC on average, whereas within herds on an individual animal level, the animals with high production still have a higher chance to get mastitis, which is in agreement with the results of (Windig et al., 2005c). This result suggests that even though the average level of SCS can be influenced by management, an antagonistic relationship on an animal level might still exist between yield and SCC, thus indicating the need to investigate the extent to which phenotypic and genetic correlations between traits can be changed by herd management.

In herds with a high average number of inseminations, DFS was shorter and FSC and NR56 were lower, compared with herds with a low average number of inseminations. Average number of inseminations had correlations with average protein production and number of animals of 0.18 and 0.14, respectively (results not shown). Herds with higher average protein production and higher number of animals also had, on average, shorter DFS and lower FSC and NR56. This trend indicates that the poorer success of first inseminations on herds with high average number of inseminations is not only explained by the stage of lactation in which insemination started, but also by the average protein production and the number of animals in the herd. Better insemination results in herds that start later with inseminating their cows, might be a result of the farmers' awareness of the poor response to early insemination, rather than better fertility per se.

Interaction of herd characteristics and breeding for health and fertility

No particular EP was superior in terms of being associated with ES for most of the traits, but ES was more often detected in situations where the EP was the average of the analyzed trait or a characteristic closely related to the trait. Although these situations might be discarded as "only scaling effects" and therefore a purely statistical issue, the change in genetic variance across environments indicates the change in importance of selection for a certain trait across environments. In situations where the EP were calculated as the average of the analyzed traits, the breeding values of the respective sires were included in the EP through the performance of their daughters. However, results from a simulation study showed that this fact does not lead to an overestimation of ES (Calus et al., 2004). Following the instruction from these authors, most EP were calculated from at least 25 animals and herd-year-season subclasses with daughters of fewer than 3 sires were deleted. This editing would thus have minimized the possible bias further.

Possible interactions between herd characteristics and breeding values for health and fertility traits were particularly notable for survival in both heifer and multiparous animals. In the heifer data, considerable changes in genetic variance were estimated for survival when herd environment was defined as fat-to-protein ratio, change in fat percentage, and change number of animals, and the genetic correlations between survival expressed in different environments were between 0.8 and 0.9. Two of these EP are indicative of the nutritional environment; lower values for change in fat percentage are associated with longer lasting and more severe negative energy balance (De Vries and Veerkamp, 2000), and fat-to-protein ratio is influenced by the amount of concentrate and fiber in the diet (Bargo et al., 2003). Therefore, it might be hypothesized that reasons for culling heifers are different in herds with different diets, because a mismatch of genetic merit of the cows and diets results in increased health and fertility problems (Veerkamp et al., 1995) such as milk fever, days to first service, and days to first heat (Pryce et al., 1999). The third environmental parameter that gave significant ES for survival in heifers was change in the number of animals. This parameter might indicate whether the farm is shrinking or expanding, and the ES might possibly be caused by different culling reasons for heifers in herds that are either expanding or shrinking. In the data of multiparous cows, the genetic correlation between survival in herds with low vs. high average age at calving was 0.65, indicating that cows in herds with a low average age at calving (and presumably an above average replacement rate) are culled for different reasons than in herds with high average age at calving. The implication of the ES for survival is demonstrated in Figure 5.1 for those sires that had most daughters in the dataset (i.e., relatively heavily used sires in the Netherlands). The sire marked with triangles had a higher EBV for survival in herds with high fat-to-protein ratio, whereas the EBV for survival of some other sires was hardly associated with the fat-to-protein ratio. These differences in patterns of the sires EBV indicate that at sire level ES has an important role, even though

parameters at a population level indicate a more limited effect of ES. Nearly all genetic correlations were above 0.8, which generally indicates that separate breeding programs for the extreme herd environments are not justified for an AI organization. However, considerable differences in ranking of top bulls across environments might occur, even at genetic correlations between environments that are above 0.9 (e.g., Powell and VanRaden, 2002). To make use of the ES and reranking of individual genotypes, or reduce the chance of a mismatch between genotype and herd environment, herd-specific breeding values might be added to customized economic selection indices (Bowman et al., 1996) for selection of bulls on a herd level. Another strategy might be to select sires which are relatively environmental insensitive, i.e., to select against ES. Application of customized selection indexes would help to fine-tune genotype and environment, whereas selection against ES actually would decrease the need to fine-tune genotype and environment.

In addition to survival, changes in genetic variance across environments were also estimated for DFS and predicted mastitis in the heifer data. Genetic variance for DFS was larger for herds with a greater change in fat percentage, a larger CIV, and a lower BCS; hence, use of sires with desirable EBV for DFS is likely to be more beneficial in those herds. Genetic variance for predicted mastitis was higher in herds with lower fat-to-protein ratio, higher number of inseminations and earlier peak date of calving; thus, use of sires with desirable EBV for mastitis is likely to be more beneficial in those herds. Changes in genetic variances of traits that are combined in a total merit index can cause reranking across environments based on the total merit index (Namkoong, 1985). The changes in genetic variances of the traits estimated on the heifer data were therefore compared with the estimated changes in genetic variances of milk, fat, and protein yield from a previous study (Calus and Veerkamp, 2003). Genetic variance of DFS increased 78% between 10th and 90th percentiles of the data based on increasing average calving interval, whereas genetic variance of milk, fat, and protein yield decreased with 14%, 8% and 15%, respectively (Calus and Veerkamp, 2003). One way to express the relative importance of selection on a trait across environments is to multiply the economic value with the environment specific genetic standard deviation of the trait. Results for yield traits and DFS combined with EP average calving interval were calculated relative to the importance of protein yield per environment (Table 5.6). The economic values were –0.08 € per kg of milk, 1 € per kg of fat, 6 € per kg of protein (NRS, 2001b) and 5.2 € per standard deviation of the fertility index (NRS, 2001a). The economic value of the fertility index was used to calculate the economic value per unit of DFS, by dividing it by the genetic standard deviation in the average environment. The relative importance of DFS compared to protein yield was twice as high in a herd with an average calving interval of 430 d, compared to a herd with an average calving interval of 370 d (Table 5.6). Hence, herd-specific breeding values in a customized index might be required to account

for changes in the weighting of traits due to different changes in genetic variances of traits across environments.

Table 5.6. *The change in relative importance to protein yield of yield traits and days to first service (DFS) across herd environments with different average calving intervals.*

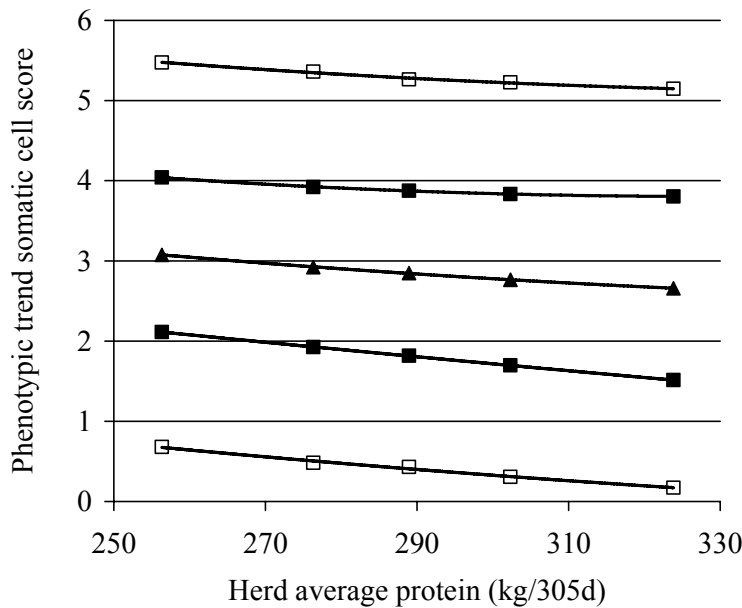| Average calving interval (d)	Trait			
	Milk	Fat	Protein	DFS
350	-0.47	0.21	1.00	0.06
370	-0.47	0.22	1.00	0.07
390	-0.46	0.22	1.00	0.09
410	-0.47	0.23	1.00	0.11
430	-0.48	0.23	1.00	0.12

Figure 5.2. *Phenotypic trend (▲) of SCS ± 2 times the genetic (■) and phenotypic standard deviation (□) across herd average protein production estimated between 10th and 90th percentiles of the cow data.*

Health and fertility risks in different herd environments

Changes in management can influence average phenotypic performance, and, as demonstrated in this study, at the same time, the genetic and phenotypic variance. An interesting question is whether these changes go hand in hand, and if improvements in mean

performance are sometimes offset by increase in variance that is so great that there is still an increase in the number of animals that perform below a threshold. In Figure 5.2, the mean phenotypic SCS ± 2 times the genetic and phenotypic standard deviation is shown as function of average protein production. With increasing average protein production, mean SCS decreased, whereas the genetic variance of SCS increased slightly (Figure 5.2). Thus, the differences in SCS between sire progeny increase when protein production per cow per year increases. The phenotypic variance however hardly increased (Figure 5.2), and, combined with the improvement in phenotypic mean, the proportion of animals that have values for SCS above any value decreased with increasing average protein production. The same was true for other situations where an improvement in phenotypic mean was accompanied by an increase in variance, suggesting an improvement in phenotypic mean usually has more impact than an increase in variance.

CONCLUSIONS

Herd-year average SCS and herd-year average CIV had a stronger association with phenotypic levels of health and fertility traits than did production level per se. Somatic cell score and predicted mastitis were lower on average in herds with higher average protein production. Genetic correlations of traits expressed in different environments were mainly close to unity, but <0.8 in a few situations. Genetic variance was generally constant across environments, but doubled across the range of some environment parameters, and the changes in genetic variance were in some situations in the opposite direction of the change in the mean of production traits. The relative importance of some fertility traits compared with yield traits doubled across environments. Both reranking of individual animals within a population and the changes in genetic variances across environments suggest that environment-specific breeding values should be estimated that could be added to customized selection indices.

Chapter 6

Estimation of genetic parameters for milk fat depression in dairy cattle

M. P. L. Calus[1,2]

M. J. Carrick[1]

R. F. Veerkamp[2]

M. E. Goddard[1,3]

[1]Department of Primary Industries Research Victoria,
Attwood, Victoria, 3049, Australia

[2]Animal Sciences Group
Division Animal Resources Development
P.O. Box 65, 8200 AB Lelystad, The Netherlands

[3]Institute of Land and Food Resources,
University of Melbourne,
Parkville, Victoria, 3052, Australia

ABSTRACT

The objective of this study was to apply reaction norm models to milk recording data to investigate genetic variation in and environmental sensitivity of susceptibility to milk fat depression (MFD). Data comprised 556,276 test-day records of 80,493 heifers in 1043 herds. Breeding values and genetic variances for fat percentage and fat yield were estimated by applying random regression models to average herd-test-day fat percentage. Genetic and permanent environmental correlations between fat yield expressed in different environments ranged, respectively, from 0.83 to 1.00 and from 0.29 to 1.00. Genetic and permanent environmental correlations between fat percentage expressed in different environments ranged, respectively, from 0.87 to 1.00 and from -0.05 to 0.99. Two traits were defined for MFD. The first trait reflected variation of milk fat percentage of animals within lactation after correction for year-season, herd-test-day, age-at-calving and stage-of-lactation. This trait had an estimated heritability of about 5% and a genetic correlation between the fifth and 95th percentile of the data of 0.50. The second trait reflected the deviation of an animal's fat percentage on a test-day from its expected fat percentage based on fat percentage on the first test-day. This trait had an estimated heritability of about 4% and a genetic correlation between the fifth and 95th percentile of the data of 0.43. The correlation between estimated breeding values of sires for the 2 MFD traits was –0.3. Our results suggest that genetic variation in susceptibility to MFD is present and that selection for reduced susceptibility to MFD is possible.

INTRODUCTION

Feeding of diets with high proportions of concentrate and low fiber to dairy cattle can result in decreased pH in the rumen (Nocek, 1997; Kennelly et al., 1999; Bargo et al., 2003) leading to depression of milk MF% (**MFD**) (Stockdale et al., 1987; Sutton, 1989; Bargo et al., 2003) and, in some cases, to (subclinical) acidosis (Nocek, 1997; Bargo et al., 2003). Although the exact mechanism is not yet known, one of the proposed theories is that milk fat synthesis is inhibited because of metabolic changes in the rumen (Griinari et al., 1998). A decrease in milk fat percentage (**MF%**) can directly lead to a financial loss if the milk price depends on MF%. In addition, it has been shown that a strong decrease in MF% in early lactation is related to a larger and longer lasting negative energy balance (De Vries and Veerkamp, 2000) and a lower first service conception (Loeffler et al., 1999). Subclinical acidosis is associated with several problems such as reduced feed intake, lower efficiency of milk production, and laminitis (Nocek, 1997). Clinical acidosis results in very sick cows.

Occurrence of MFD can be observed by monitoring changes in MF% or fat-to-protein ratio, as protein percentage is reported to be either unchanged or enhanced during MFD (Bargo et al., 2003). Additional symptoms for subclinical acidosis are reduced feed intake, BW loss, diarrhea, and lameness (Nocek, 1997). Gröhn et al. (1989) reported that acidosis

mostly occurs in early lactation. Prevention of MFD and acidosis can be achieved by feeding a balanced diet or by adding a buffer to the diet (Kennelly et al., 1999). Feeding balanced diets might be impossible in pasture-based systems where diets depend on seasonal supply of grass and where concentrate is used to compensate for temporarily low availability of grass rather than as a permanent component of the diet. In situations where concentrates are used as a structural supplement to pasture, occurrence of MFD is still reported (Bargo et al., 2002; Bargo et al., 2003) indicating that not only temporarily, but also continuously high proportions of concentrate in the diet can cause MFD.

Although most experiments that challenged animals to express MFD reported a depression in MF%, some also reported a depression of fat yield (Gaynor et al., 1994; Griinari et al., 1998), and others did not find an effect on fat yield (Bargo et al., 2002). This is in agreement with the general observation that diets with a high proportion of concentrates or low fiber may cause depressed MF%, while milk yield is either unchanged or enhanced (Bargo et al., 2003). A question from a breeder's point of view is whether susceptibility to MFD is correlated to breeding values for fat yield and MF%, and also to breeding values for health and fertility.

Although much is known about causal relations between composition of the diet and MFD, little is known about differences in MFD between genotypes (i.e., genetic variance of MFD) or differences in responses of genotypes to different diets (i.e., genotype × environment interaction of MFD). To estimate genetic variance and genotype × environment interaction, a large number of records are needed, whereas MFD is typically examined in feeding experiments, where responses are measured within small treatment groups (Stockdale et al., 1987; Bargo et al., 2002). However, as occurrence of MFD can be observed as a change in MF%, it should be possible to define MFD on an individual and an environmental level using test-day records for MF%. This practice would enable the use of large data sets available from milk recording to investigate occurrence of MFD on an environmental and a genetic level.

Random regression models (**RRM**) modeling reaction norms (Veerkamp and Goddard, 1998; Kolmodin et al., 2002; Calus and Veerkamp, 2003; Hayes et al., 2003) can be used to investigate environmental sensitivity of EBV and genetic variances. In an RRM, an animal's breeding value for a particular trait is modeled as a function of a particular environmental variable. Thus, we need to choose a trait measured on each individual cow and an environmental indicator variable. Because MFD is a depression in MF%, it seems logical to choose MF% as the trait and average herd-test-day MF% (**AHTDF%**) as the environmental indicator. Then a cow that was highly susceptible to MFD would be expected to have a very low MF% on days when AHTDF% was low and, perhaps, a normal MF% on days when AHTDF% was high. However, other factors, such as management, stage of lactation, or frequency of milking, affect MF% as well as MFD. For instance, low MF% could reflect an increase in milk volume due to good nutrition (Bargo et al., 2003), or high MF% a decrease in milk volume due to mastitis (Windig et al., 2005b). Therefore, we have tested a range of

response traits to determine which might be the best at identifying bulls whose daughters are susceptible to MFD. In addition to the use of MF% and fat yield as response traits of MFD, we investigated other indicators based on the pattern of MF% across test-days within lactation.

The objective of this study was to investigate genetic variation in and environmental sensitivity of susceptibility to MFD, affecting MF% and fat yield, from milk recording data using reaction norm models.

MATERIALS AND METHODS
Data

The original data set was extracted from the Australian Dairy Herd Improvement Scheme. Initial edits of the data are described by Hayes et al. (2003) and included deleting herd-test-days with fewer than 20 heifers. Additionally, herd-test-days with average MF% above 5.5% were deleted. The final data set for MF% and fat yield contained 16,344 herd-test-days comprising 556,276 test-day records of 80,493 Holstein Friesian heifers in 1043 herds. The final data set for the MFD trait defined on a lactation level (**MFDLAC**; see below) contained 3032 herd-year-season subclasses comprising 68,907 Holstein Friesian heifers in 960 herds, after deleting heifers with fewer than 5 test-day records, herd-year-season subclasses with fewer than 5 heifers, and sires with fewer than 5 daughters in the data. The final data set for the MFD trait defined on a test-day level (**MFDTD**; see below) contained 14,241 herd-test-days comprising 484,219 records of 78,256 Holstein-Friesian heifers in 1021 herds, after deleting herd-test-days with <5 heifers and deleting sires with <5 daughters in the data. The relationship matrix contained 6170 animals, of which 1539 were sires with daughters in the data.

Definition of traits and environment

Traits for MFD. As well as fat yield and MF%, 2 other traits were derived from the pattern of MF% across test-days. The lactation curve for MF% of an animal affected by MFD will fall below the normal lactation curve. Based on deviation of a lactation curve for MF% from a "nonaffected" lactation curve, MFD can be divided into 2 dimensions: magnitude and duration. Two traits to reflect MFD were defined: one on a lactation level (MFDLAC) and one on a test-day level (MFDTD). The MFDLAC trait was defined as the standard deviation of residual test-day MF% of a heifer. Residual MF% were calculated using a model with fixed effects correcting for year-season, herd-test-day, and a third-order fixed regression on age at calving and an eighth-order fixed regression on DIM. This model yielded residuals for MF% for each test-day record. The MFDLAC was calculated for each heifer with at least 6 test-days as the standard deviation of residual MF% across test-days. Animals with high susceptibility to MFD are more likely to have one or more test-days with reduced MF%, resulting in a

higher variation of milk MF% during the lactation and therefore a higher value for MFDLAC than animals with low susceptibility to MFD.

Dalley (2002) defined MFD on a herd level as a decline in MF% of more than 0.4% in 10 days. Following Dalley (2002), we adopted the minimum value of the decline of 0.4% to define MFD on an individual level. The MFDTD trait was defined for all heifers with at least 2 test-day records and was based on the difference between MF% on the first test-day and all subsequent test-days. Individual test-day records for MF% were corrected for age at calving and DIM using smoothing splines in ASREML (Gilmour et al., 2002b). From all second and later test-day records for MF% of each heifer, the MF% from its first test-day was subtracted. Based on the obtained difference on test-day i (Δfat%$_{t=i,t=1}$), MFDTD was defined as follows:

$$\text{If } \Delta\text{fat\%}_{t=i,t=1} < -0.4\%$$
$$\text{Then: MFDTD} = \Delta\text{fat\%}_{t=i,t=1} + 0.4$$
$$\text{Otherwise: MFDTD} = 0.$$

This resulted in a continuously distributed trait with a maximum value of 0. Animals with high susceptibility to MFD were expected to have a lower value for MFDTD than animals with low susceptibility to MFD.

Definition of environment. Individual test-day milk, fat and protein yields, corrected for age at calving and DIM using cubic splines in ASREML (Gilmour et al., 2002b), were used to calculate herd-test-day averages for fat yield, MF% and fat-to-protein ratio. The obtained herd-test-day averages were used to calculate differences between pairs of consecutive herd-test-days that were less than 70 d apart. Correlations among averages of herd-test-days were calculated to determine relationships among them.

Estimation of variance components for fat percentage and fat yield

Variance components for fat yield and MF% were estimated using a RRM. Variance components for fat yield were also estimated using a bivariate repeatability model (**BVM**), to compare results of the RRM. Therefore, the data were divided in 10 subsets based on increasing AHTDF%, and the BVM was applied to all possible pairs of subsets of the data. The RRM contained the same effects as the one applied by Hayes et al. (2003), although heterogeneous residual variances were included in our model. The BVM and RRM included fixed effects for mean of the trait, herd-test-day, year-season subclasses, and fixed polynomial regressions for age at calving and DIM. A fixed polynomial regression was applied to DIM with an arbitrary order of 8 to account for the average lactation curve. The BVM included correlated random effects for sire and cow in both environments, to account for genetic and permanent environment effects respectively. The RRM included random regressions for both

sire and cow on AHTDF%. In both the BVM and the RRM, a separate residual variance was estimated for each subset of the data. The general model was:

$$Y_{ijklm} = \mu + HTD_i + YS_j + \sum_{n=1}^{3} A_n x_n + \sum_{o=1}^{8} D_o Z_o + sire_k + cow_l + E_{ijklm}$$

where Y_{ijklm} is the phenotypic performance of heifer l, μ is the average performance over all animals, HTD_i is a fixed effect for herd-test-day i, YS_j is a fixed effect for year season (defined from January to June and July to December for each year) j, A_n is coefficient n of a third order fixed regression on age at calving, x_n is the nth order polynomial corresponding to age at calving, D_o is coefficient o of an eighth order fixed regression on DIM, Z_o is the oth order polynomial corresponding to DIM, $sire_k$ is a random sire effect for sire k, cow_l is a random within sire genetic plus permanent environmental effect for heifer l, E_{ijklm} is the residual effect of heifer l in group of environments m ($m = 1,2,…,10$).

In the BVM, all fixed effects other than HTD and all random effects were estimated for both environments. In the RRM, all fixed effects were fitted for all environments together. Sire and cow effects were fitted as random regressions on orthogonal polynomials. We chose Legendre polynomials following the practice of Kirkpatrick et al. (1990). The orders of the random regressions for sire and cow effects were assumed equal and were increased until the highest significant order was reached based on the likelihood ratio test. In both the BVM and RRM, the residual covariance between environments was assumed to be zero.

Estimation of variance components for MFD traits

The trait MFDTD was analyzed with the RRM as described for MF% and fat yield, with random regressions for sire and cow on AHTDF%. For the trait MFDLAC, a comparable RRM was applied. This trait was however defined on a lactation level and therefore the model was simplified to:

$$Y_{iklm} = \mu + HYS_i + \sum_{n=1}^{3} A_n x_n + sire_k + E_{iklm}$$

where Y_{iklm} is the phenotypic performance for MFDLAC of heifer l, μ is the average performance over all animals, HYS_i is a fixed effect for herd-year-season subclass (defined from January till June and July till December for each herd-year) i, A_n is coefficient n of a third order fixed regression on age at calving, x_n is the nth order polynomial corresponding to age at calving, $sire_k$ is a random sire effect for sire k, E_{iklm} is the residual effect of heifer l in group of environments m ($m = 1,2,…,5$).

The random effect for sire, ($sire_k$) was a random regression on an average value for AHTDF% during the lactation of an animal. This average was calculated for each heifer as the average of the values for AHTDF% of the test-days during which the heifer had records.

Calculation of parameters from estimated variance components in different environments

Both models estimated variances and covariances for cow, reflecting the within-sire genetic plus permanent environmental (co)variance. Estimated permanent environmental (co)variances for both models were obtained by subtracting 3 times the estimated sire (co)variance from the estimated (co)variance for cow.

The BVM was applied to all 45 pairs of subsets of the data for fat yield, to estimate genetic and permanent environmental correlations between fat yield expressed in all subsets of the data. Hence, 9 estimates for each estimated variance component and its standard error were obtained and averaged to get final estimates. The RRM estimated sire and cow (co)variances as function of the environment, which were used to derive estimated genetic and permanent environmental (co)variances of the RRM.

The heritability in an environment was calculated as 4 times the estimated sire variance, divided by the sum of the sire, cow, and residual variance. For the trait MFDLAC, the heritability in an environment was calculated as 4 times the estimated sire variance, divided by the sum of the sire and residual variance. For the trait MFDTD, the repeatability in an environment was calculated as the sum of the sire and cow variance, divided by the sum of the sire, cow and residual variance.

RESULTS

Phenotypic levels of MFD

Mean, standard deviation, and the ranges of herd-test-day averages and differences between averages of consecutive herd-test-days for fat yield, MF% and fat-to-protein ratio are shown in Table 6.1. The average herd-test-day MF% ranged from 2.18 to 5.48% and the change in average MF% from one to the next herd-test-day ranged from −1.82 to 1.53%. Correlations between herd-test-day averages, calculated from 16,344 herd-test-days, were 0.03 (SE = 0.01) between fat yield and MF%, -0.12 (SE = 0.01) between fat yield and fat-to-protein ratio, and 0.82 (SE = 0.00) between average MF% and fat-to-protein ratio.

The value for the heifers for the trait MFDLAC ranged from 0.00 to 1.43, with an average of 0.28. The average value for MFDLAC within herd-year-season subclasses ranged from 0.03 to 0.31. The average value for MFDLAC within herd ranged from 0.04 to 0.21.

The values for MFDTD ranged from −3.4 to 0%, with an average of -0.1%. Based on the trait MFDTD, 44.9% of the heifers had one or more affected test-days resulting in 22.1% of all records being affected by MFD. In 93.7% of the herd-test-days one or more affected heifers were identified. The percentage of affected heifers based on MFDTD decreased with increasing days in milking (Figure 6.1).

Table 6.1. *Mean, standard deviation, and range of herd-test-day averages and differences between averages of pairs of consecutive herd-test-days (Δ) for fat yield, fat percentage, and fat-to-protein ratio.*

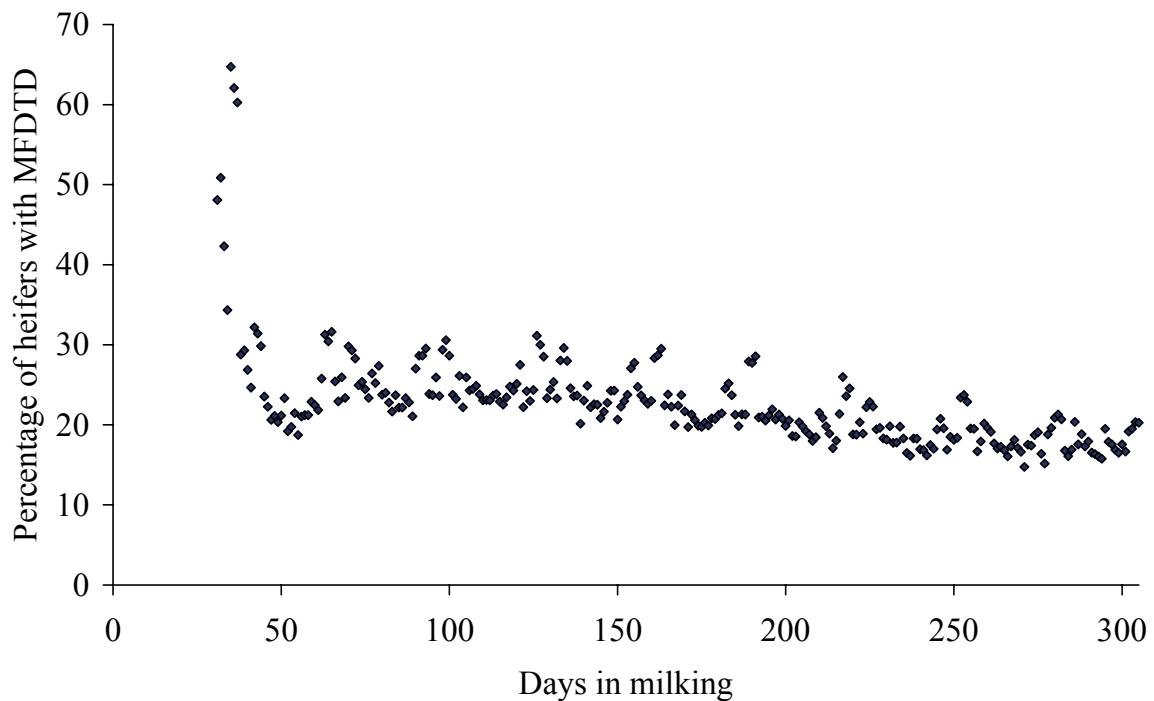| Trait	Mean	Standard deviation	Minimum	Maximum
Fat (kg)	0.77	0.13	0.37	1.21
Fat percentage	3.87	0.33	2.18	5.48
Fat-to-protein ratio	1.21	0.11	0.66	1.76
ΔFat (kg)	-0.01	0.08	-0.42	0.33
ΔFat percentage	0.02	0.28	-1.82	1.53
ΔFat-to-protein ratio	0.01	0.09	-0.65	0.58

Figure 6.1. *The percentage of heifers showing milk fat depression based on test-days (MFDTD), across days in milking.*

Estimated variance components for fat yield

Estimated genetic variances based on 556,276 test-day records for the BVM and the RRM with environment defined as AHTDF% are shown in Figure 6.2. Estimated genetic variances from the BVM are given with a range of their standard errors (Figure 6.2). The quadratic RRM was the highest significant order (a cubic order did not converge). In the range of environments where most animals were situated (i.e., fifth percentile of the data was 3.23% and 95th percentile was 4.43%), the sire variance of the linear and quadratic RRM was almost the same. Estimated heritabilities and genetic and permanent environmental correlations
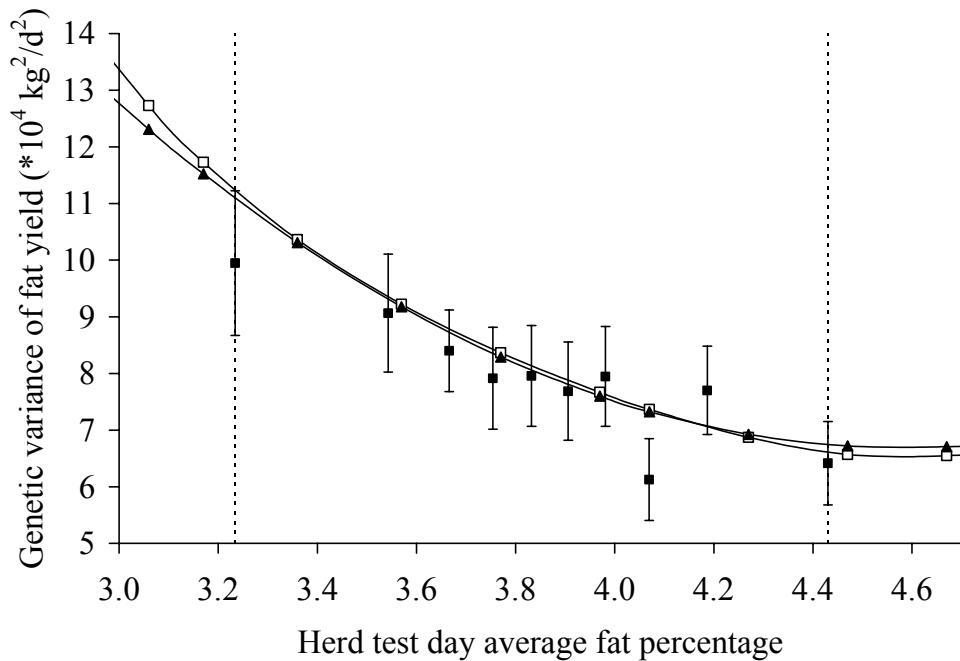
Figure 6.2. *Genetic variances of fat yield ($\times 10^4$ kg^2/d^2), with environment defined as herd-test-day average fat percentage, estimated with a linear random regression model (RRM) (▲), a quadratic RRM (□), and for the different subsets of the data estimated with a bivariate repeatability model (■). Fifth and 95th percentiles of the data are shown as dotted lines.*

between the different subsets of the data estimated with the BVM (results not shown) and quadratic RRM (Table 6.2) were comparable. The average AHTDF% of each subset of the data was used as an environmental variable to calculate those parameters. Estimated heritabilities ranged, for both models, from 0.17 to 0.20 (SE 0.02 to 0.03 for the BVM and 0.02 for the RRM). Estimated genetic correlations of fat yield between subsets of the data ranged for both models from 0.83 to 1.00 (SE 0.00 to 0.05 for the BVM and 0.00 to 0.04 for the RRM). Permanent environmental correlations ranged from 0.51 to 1.00 (SE 0.01 to 0.04) for the BVM and from 0.29 to 1.00 for the RRM (SE 0.00 to 0.01). Standard errors were higher for the BVM than the RRM in all situations.

The EBV for fat yield estimated with the quadratic random regression on AHTDF% of 10 sires with most daughters in the data are shown in Figure 6.3. These 10 sires all had more than 9000 daughters in the data. The EBV are shown for a range of environments with an AHTDF% between the fifth and 95th percentiles of the data. The course of EBV for fat yield across AHTDF% differed among those 10 sires (Figure 6.3), resulting in some reranking of sires across environments.

Table 6.2. *Estimates from the quadratic random regression model for heritabilities[1] of fat yield (on the diagonal) in 10 subsets of the data defined on increasing average herd-test-day fat percentage (AHTDF%), and genetic correlations[2] (above diagonal), and permanent environmental correlations[3] (under diagonal) between fat yield expressed in the different subsets of the data.*

AHTDF%	3.23	3.54	3.67	3.75	3.83	3.91	3.98	4.07	4.19	4.43
3.23	0.19	0.99	0.97	0.96	0.95	0.94	0.93	0.91	0.89	0.83
3.54	0.96	0.20	1.00	0.99	0.99	0.98	0.98	0.97	0.95	0.90
3.67	0.92	0.99	0.19	1.00	1.00	0.99	0.99	0.98	0.96	0.92
3.75	0.88	0.98	1.00	0.20	1.00	1.00	0.99	0.99	0.97	0.93
3.83	0.84	0.95	0.98	1.00	0.20	1.00	1.00	0.99	0.98	0.94
3.91	0.79	0.92	0.96	0.99	1.00	0.19	1.00	1.00	0.99	0.95
3.98	0.73	0.89	0.94	0.97	0.98	1.00	0.20	1.00	0.99	0.96
4.07	0.66	0.83	0.90	0.93	0.96	0.98	0.99	0.19	1.00	0.98
4.19	0.54	0.75	0.82	0.87	0.91	0.94	0.97	0.99	0.18	0.99
4.43	0.29	0.52	0.62	0.69	0.75	0.80	0.85	0.90	0.96	0.17

[1]SE of the heritabilities were 0.02.

[2]SE of the genetic correlations ranged from 0.00 to 0.04.

[3]SE of the permanent environmental correlations ranged from 0.00 to 0.01.

Estimated variance components for fat percentage

The linear RRM was the highest significant order for MF%. The sire variance for MF% as a function of AHTDF% estimated with a linear RRM is shown in Figure 6.4. Estimated heritabilities and genetic and permanent environmental correlations between the different subsets of the data are shown in Table 6.3. The estimated heritability for MF% ranged from 0.40 to 0.53 (SE = 0.03) between the fifth and 95th percentile of the data. The estimated genetic and permanent environmental correlations between MF% expressed in the fifth and 95th percentile of the data were 0.87 (SE = 0.02) and -0.05 (SE = 0.09), respectively.
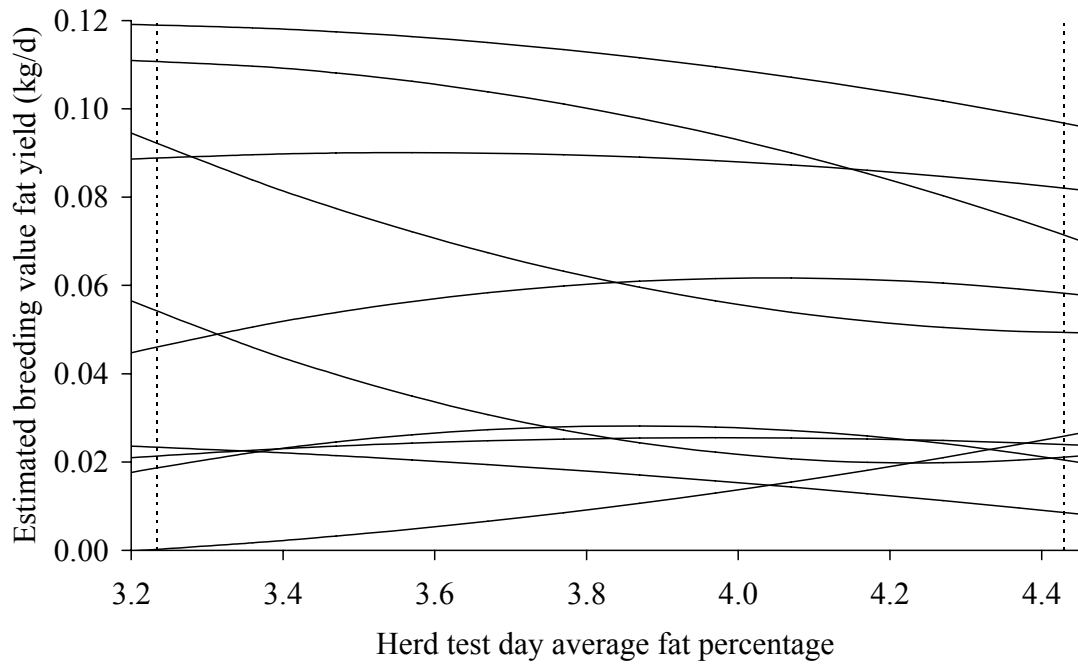
Figure 6.3. *Sire breeding values for fat yield estimated with a quadratic random regression on average herd-test-day fat percentage, of 10 sires with most daughters in the data. Fifth and 95th percentiles of the data are shown as dotted lines.*
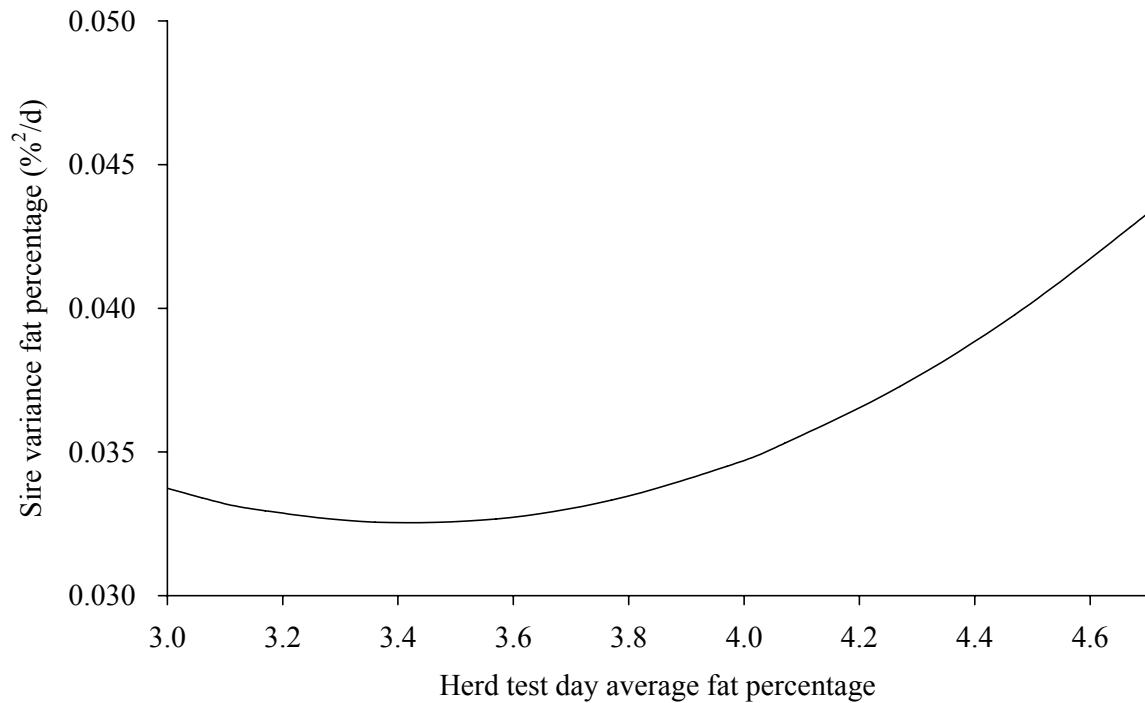

Figure 6.4. *Sire variance ($\%^2/d$) for fat percentage estimated with a linear random regression on average herd-test-day fat percentage.*

Table 6.3. *Estimates from the linear random regression model for heritabilities[1] of fat percentage (on the diagonal) in 10 subsets of the data defined on increasing average herd-test-day fat percentage (AHTDF%), and genetic correlations[2] (above diagonal) and permanent environmental correlations[3] (under diagonal) between fat percentage expressed in the different subsets of the data.*

AHTDF%	3.23	3.54	3.67	3.75	3.83	3.91	3.98	4.07	4.19	4.43
3.23	0.40	0.99	0.98	0.97	0.96	0.96	0.95	0.93	0.91	0.87
3.54	0.93	0.48	1.00	1.00	0.99	0.99	0.98	0.97	0.96	0.93
3.67	0.84	0.98	0.50	1.00	1.00	0.99	0.99	0.98	0.98	0.95
3.75	0.75	0.94	0.99	0.51	1.00	1.00	1.00	0.99	0.98	0.96
3.83	0.65	0.88	0.96	0.99	0.52	1.00	1.00	0.99	0.99	0.97
3.91	0.55	0.82	0.91	0.97	0.99	0.52	1.00	1.00	0.99	0.98
3.98	0.44	0.74	0.86	0.93	0.97	0.99	0.52	1.00	1.00	0.98
4.07	0.32	0.65	0.78	0.87	0.93	0.97	0.99	0.51	1.00	0.99
4.19	0.18	0.53	0.68	0.79	0.86	0.92	0.96	0.99	0.49	1.00
4.43	-0.05	0.32	0.50	0.63	0.73	0.81	0.87	0.93	0.97	0.47

[1]SE of the heritabilities were 0.03.

[2]SE of the genetic correlations ranged from 0.00 to 0.02.

[3]SE of the permanent environmental correlations ranged from 0.00 to 0.09.

Estimated variance components for MFDLAC and MFDTD

The linear RRM was the highest significant order for MFDLAC. The sire variance for the linear RRM is shown in Figure 6.5 as a function of average AHTDF%. The heritability of MFDLAC was 0.056 (SE = 0.015) in both the fifth and 95th percentile of the data and had a value of 0.046 (SE = 0.008) in the average environment. The genetic correlation between MFDLAC expressed in the fifth and 95th percentile of the data was 0.50 (SE = 0.20).

The linear RRM was the highest significant order for MFDTD (a quadratic order did not converge). The residual variance ranged from 0.056 to $0.008\%^2$ in environments with increasing AHTDF%. The sire variance is shown in Figure 6.6 as a function of average AHTDF%. The permanent environmental variance across environments was 5 to 11 times as high as the sire variance. The heritability of MFDTD was 0.046 (SE = 0.006) in the fifth percentile and 0.018 (SE = 0.006) in the 95th percentile of the data and had a maximum value of 0.053 (SE = 0.009), within this range of environments. The repeatability of MFDTD had values between 0.53 and 0.75 (SE = 0.00) in the range of the data between the fifth and 95th percentiles. The genetic and permanent environmental correlations between MFDTD expressed in the fifth and 95th percentile of the data were respectively 0.43 (SE = 0.15) and 0.65 (SE = 0.01).
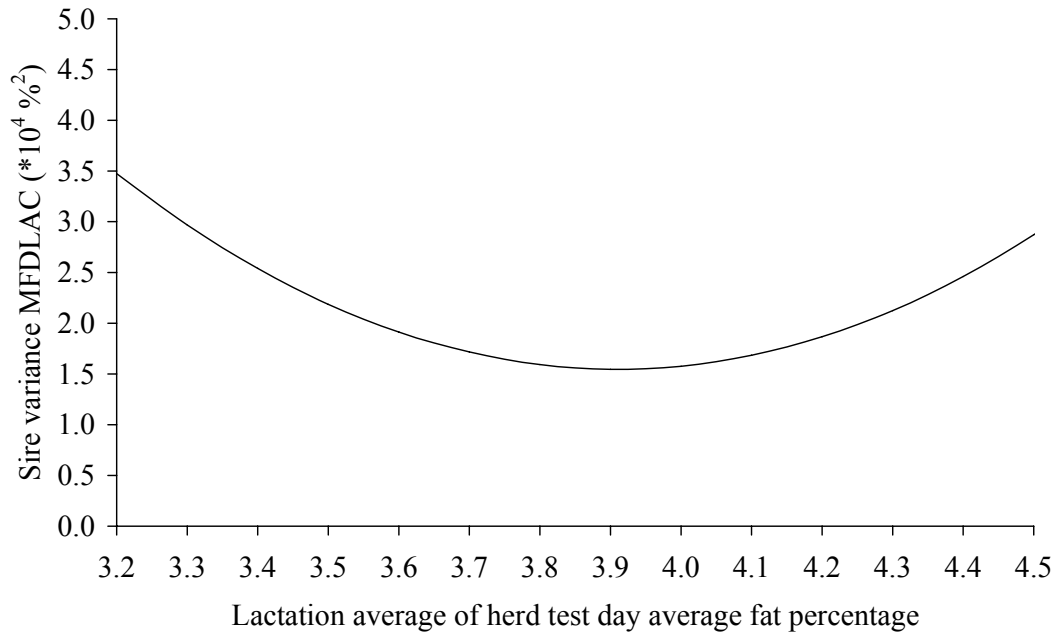
Figure 6.5. *Sire variance ($\times 10^4$ $\%^2$) for milk fat depression defined on a lactation level (MFDLAC) estimated with a linear random regression on lactation average of average herd-test-day fat percentage.*
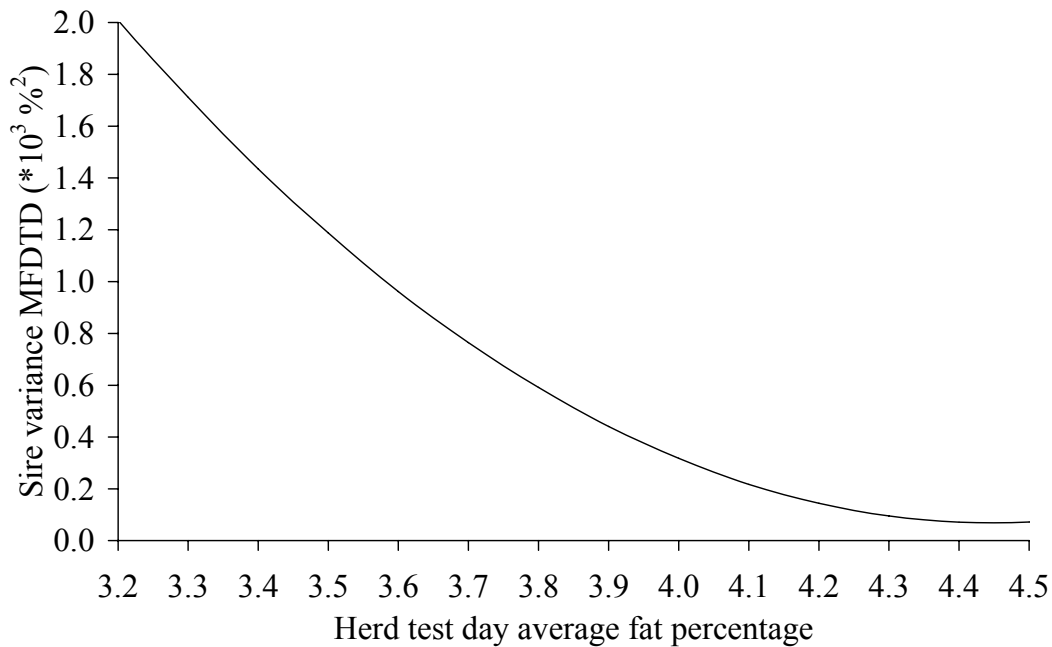

Figure 6.6. *Sire variance ($\times 10^3$ $\%^2$) for milk fat depression defined on a test-day level (MFDTD) estimated with a linear random regression on average herd-test-day fat percentage.*

Table 6.4. *Correlations between estimated sires breeding values calculated for environments with average herd-test-day fat percentage (AHTDF%) of 3.3% and 4.4%, respectively for milk fat depression defined on a lactation level (MFDLAC) or a test-day level (MFDTD), fat percentage (Fat%), fat yield (Fat), and differences of estimated breeding values for fat percentage (ΔFat%) and fat yield (ΔFat) between the 2 environments (SE ranged from 0.01 to 0.03).*

AHTDF%	Trait	MFDTD	Fat%	Fat	ΔFat%	ΔFat
3.3%	MFDLAC	-0.28	0.09	-0.07	0.30	0.13
	MFDTD		0.33	0.21	-0.22	-0.42
	Fat%			0.51	0.07	-0.48
	Fat				-0.08	-0.72
4.4%	MFDLAC	-0.34	-0.08	-0.18	0.22	0.24
	MFDTD		0.28	0.00	0.17	-0.28
	Fat%			0.40	0.40	-0.35
	Fat				0.03	-0.45
	ΔFat%					0.27

Correlations between estimated breeding values for the different traits

The EBV for sires with daughters in the data were calculated for the traits MFDLAC, MFDTD, MF%, and fat yield for environments with values of 3.3 and 4.4% for AHTDF%.

Additionally, the changes in EBV for MF% and fat yield between both environments were calculated as a measure of the environmental sensitivity of the EBV of the sire for MF% and fat yield. These changes were calculated as EBV at AHTDF% = 4.4% minus EBV at AHTDF% = 3.3%. Correlations between all obtained EBV and changes between EBV are shown in Table 6.4. The chosen values for AHTDF% represented approximately the fifth and 95th percentiles of the data for the 4 traits. The EBV for MFDLAC and MFDTD were moderately correlated. The EBV for MFDLAC had small correlations with EBV for MF% and fat yield. The EBV for MFDTD had moderate positive correlations with EBV for MF%. The EBV for MFDTD had a positive correlation with EBV for fat yield at AHTDF% of 3.3%, but a zero correlation at AHTDF% of 4.4%.

The EBV for MFDLAC were positively correlated to changes in EBV for MF% and fat yield. The EBV for MFDTD were negatively correlated to changes in EBV for MF% and fat yield, except for the combination of MFDTD at AHTDF% of 4.4% and change in EBV for MF%.

The EBV for MF% and fat yield were hardly correlated to change in EBV for MF%, except for the combination of MF% and change in fat at AHTDF% of 4.4%. The EBV for MF% and fat yield were negatively correlated to change in EBV for fat yield. Changes in EBV for MF% and fat yield were positively correlated.

DISCUSSION

Definition of environment

Both MF% and fat-to-protein ratio can be used as indicators for MFD (Bargo et al., 2003). Correlations between herd-test-day average MF% and fat-to-protein ratio were higher than 0.8, indicating that use of either MF% or fat-to-protein ratio as a definition of the environment would yield similar results.

Next to AHTDF%, the change of AHTDF% between consecutive test-days was used as definition of environment (results not shown) in the analysis for fat yield and MF%. Defining the environment by change in AHTDF% led to less reranking of sires than did defining the environment by AHTDF%. Change in AHTDF% was assigned to the latter of each pair of consecutive herd-test-days that were maximum 70 d apart. Hence, records of the first herd-test-day in seasonal calving herds had no value for change in AHTDF%. As it is reported that more than a third of the cases of acidosis occurs in the first month after calving (Gröhn and Bruss, 1990), use of change in AHTDF% as definition of environment might have resulted in losing relatively a lot of records in early lactation that were subject to MFD.

Environmental sensitivity of fat yield

Preliminary analyses indicated highest genetic variance of fat yield in environments with low AHTDF%. The correlation between herd-test-day average fat yield and AHTDF% was 0.03, indicating that differences in genetic variance for fat yield across different levels of AHTDF% could not be explained by the average fat yield for a given herd-test-day. As limited information was available in extreme environments, the BVM was applied to compare results. The application of a multivariate model including all subsets of the data at once would have been theoretically more appealing, but computationally challenging and given the number of traits was likely to result in nonpositive definite variance matrices (Hill and Thompson, 1978). The BVM included only 2 out of 10 subsets of the data at once and, therefore, estimates for genetic and permanent environmental correlations between fat yield expressed in different environments were calculated by using only 20% of the data in the BVM. This is likely to be the reason that, in general, the standard errors of heritabilities and genetic and permanent environmental correlations of fat yield expressed in different environments were lower for the RRM than for the BVM.

The results showed that between the fifth and 95th percentile of the data, the genetic variance for fat yield increased with decreasing AHTDF%. This increase in genetic variance could be because susceptibility to MFD adds to the variation when AHTDF% is low. Although there was limited reranking of sires across environments for their EBV of fat yield, the genetic correlations for fat yield expressed in different environments did indicate that fat yield is a different trait in about 10% of the environments with lowest AHTDF%. As argued by Hayes et al. (2003), the composition of milk is influenced by DIM which could lead to

confounding between DIM and AHTDF% in seasonal calving herds. That is, it was possible that the reranking of sires for fat was actually due to differences in shape of the lactation curve and not to differences in susceptibility to MFD. To exclude this possibility, following Hayes et al. (2003), the correlation between AHTDF% and DIM was estimated and a RRM with quadratic random regressions for sire and cow on both AHTDF% and DIM was fitted. The AHTDF% had a correlation with DIM of –0.15. An RRM with quadratic random regressions on DIM and AHTDF% yielded estimated genetic correlations between fat yield expressed in the subsets of the data with the lowest and highest AHTDF% ranging from 0.85 to 0.90, across DIM. The corresponding genetic correlation of the RRM with a quadratic random regression on AHTDF% had a value of 0.83 (Table 6.2). This suggests that the original estimate of 0.83 is, at most, minimally increased when genetic variation in shape of lactation curve is included in the model.

Environmental sensitivity of milk fat percentage

The estimated genetic correlation between the fifth and the 95th percentile of the data was higher for MF% than for fat yield. The estimated permanent environmental correlation between the fifth and the 95th percentile of the data was, however, lower for MF% than fat yield, i.e., –0.05 vs. 0.29. This indicates that both MF% and fat yield are controlled by different permanent environmental factors in environments with low vs. high AHTDF%. In other words, cows vary widely for nongenetic reasons in their susceptibility to MFD. The difference in permanent environmental correlations for MF% and fat yield is likely a result of different nongenetic factors acting on milk yield in environments with low vs. high AHTDF%.

Milk fat depression traits

For MFDTD, 2 test-day records were sufficient to calculate at least one record for an animal, whereas a minimum of 5 test-day records per animal was required for MFDLAC. This explains the difference in number of heifers included in the analysis for MFDLAC and MFDTD. Milk fat percentage was corrected for stage of lactation and age at calving in the calculations of MFDLAC and MFDTD. However, other factors influencing MF%, such as incomplete milking, could possibly lead to false positive records for MFD. The definitions of MFDLAC and MFDTD imply that there is no individual variation in the shape of the lactation curve for MF%, and that certain decreases (or changes) in MF% have the same importance regardless the stage of lactation or MF% before the change. In the case that an animal has MFD on all test-days within a lactation resulting in comparable reduction of its MF% on all test-days, both definitions would fail to identify that she has MFD. The current literature provides no indication how the definitions for MFD based on milk recording data could be refined to overcome these limitations.

Both MFD traits tried to capture magnitude and duration of MFD. Arguably, the magnitude of MFD could be thought of as the susceptibility to MFD, while the duration could be thought of as the ability to recover from MFD. Animal experiments in which MFD is studied usually reflect the transition to a diet with higher proportions of concentrate and lower proportions of fibre, whereas recovery from MFD typically would be driven by transition to a diet with lower proportions of concentrate and higher proportions of fiber. With regard to this hypothesis, an important question would be to determine if susceptibility to and ability to recover from MFD are strongly correlated. The answer to this question would give more insight into whether MFD can simply be defined on a lactation level, without distinction between magnitude and duration of MFD, or that MFD should be defined on a test-day level.

In the average environment, i.e., an environment with an AHTDF% of 3.87%, MFDLAC and MFDTD had heritabilities of 0.045 and 0.042, respectively. The reliability of an EBV is $r^2 = n_e/(n_e + \lambda)$, where n_e is the effective daughter size and $\lambda = (4 - h^2)/h^2$ for a sire model. To obtain a reliability of 60%, 132 effective daughters for MFDLAC and 141 for MFDTD were required for the data set used in our analyses. These figures indicate that active selection based on a trait reflecting MFD is possible.

Correlation between EBV for the different traits

High susceptibility to MFD should be indicated by low values of MF% and fat yield at AHTDF% of 3.3%, high values for MFDLAC, low values of MFDTD at AHTDF% of 3.3% and perhaps AHTDF% of 4.4%, and low values of change in MF% and fat yield.

Correlations among EBV were nearly all in the expected direction but not high. Correlations that were not in the expected direction were correlations of fat yield and change in MF% with MFDTD at AHTDF% of 4.4%, which agrees with the idea that MFD plays virtually no role in environments with high AHTDF% and therefore will hardly influence fat yield and MF% in those environments. One combination with a moderate correlation (of 0.42) is change in fat yield and MFDTD at AHTDF% of 3.3%. These are perhaps the best candidates for identifying sires whose daughters are susceptible to MFD. However, before either could be recommended, more evidence is required to support the hypothesis that they identify sires whose daughters are prone to some important outcome such as clinical acidosis or laminitis.

Changes in milk yield and composition between test-days can represent a healthy response to increased nutrition rather than a pathological response such a MFD. One of the challenges in evaluating sires is to distinguish between these 2 possibilities. For this reason, we prefer the use of fat yield to MF% as a dependent variable because an increase in milk volume causing a decrease in MF% is not pathological, but a decrease in fat yield due to increased grain feeding may be.

CONCLUSIONS

Genetic variation exists between animals in susceptibility to MFD. Data recorded routinely by milk recording agencies can be used to evaluate sires for this trait by using a random regression analysis. Selection for low susceptibility could be based on sires whose daughters show little decline in fat yield between test-days with high and low AHTDF% or sires whose daughters show a high value of MFDTD when the herd AHTDF% is low. However, EBV for MFD susceptibility calculated by different methods are not highly correlated and evidence for a genetic correlation with clinical symptoms of MFD would be necessary before advocating the use of any EBV for susceptibility to MFD.

Chapter 7

Genotype by environment interaction of somatic cell score across bulk milk somatic cell count and days in milk

M. P. L. Calus[1]

L. L. G. Janss[1,2]

R. F. Veerkamp[1]

[1]Animal Sciences Group
P.O. Box 65, 8200 AB Lelystad, The Netherlands

[2]Statistical Animal Genetics Group, Institute of Animal Science
Swiss Federal Institute of Technology, ETH Zentrum,
CH 8092 Zurich, Switzerland

ABSTRACT

The objective of this paper was to investigate the importance of genotype × environment interaction (G×E) for somatic cell score (SCS) across bulk milk somatic cell count (BMSCC), days in milk (DIM) and their interaction. Variance components were estimated with a model including random regressions for each sire on herd test-day BMSCC, DIM, and the interaction of BMSCC and DIM. The analyzed data set contained 344,029 test-day records of 24,125 cows, sired by 182 bulls, in 461 herds comprising 13,563 herd test-days. In early lactation, considerable G×E effects were detected for SCS, indicated by threefold higher genetic variance for SCS at high BMSCC compared with SCS at low BMSCC, and a genetic correlation of 0.72 between SCS at low and SCS at high BMSCC. Estimated G×E effects were smaller during late lactation. Genetic correlations between SCS at the same level of BMSCC, across DIM, were between 0.43 and 0.65. The lowest genetic correlation between SCS measures on any two possible combinations of BMSCC and DIM was 0.42. Correlated responses in SCS across BMSCC and DIM were in some occasions less than half the direct response to selection in the response environment. Responses to selection were reasonably high among environments in the second half of the lactation, while responses to selection between environments early and late in lactation tended to be low. Selection for reduced SCS yielded the highest direct response early in lactation at high BMSCC.

INTRODUCTION

Individual measures of SCC of dairy cows are used as an indicator trait for mastitis. Management and breeding decisions aim to reduce SCC, as a way to decrease the incidence of mastitis (Emanuelson, 1988; Weller et al., 1992; Philipsson et al., 1995). Another reason to reduce SCC, is to decrease bulk milk somatic cell count (**BMSCC**), as a BMSCC above a certain value results in a discount in milk price for the farmer (Veerkamp et al., 1998; Productschap Zuivel, 2004). A wide range of BMSCC is present across herds, which is at least partly explained by differences in management between herds (Barkema et al., 1998a; Barkema et al., 1999b), and is also related to presence of mastitis pathogens, since an important part of the genetic variances in SCC is caused by mastitis. Therefore, the question rises whether these management differences reflected in BMSCC affect genetic parameters for SCC and responses to selection for reduced SCC.

Selection responses could be affected if genotype × environment interaction (**G×E**), also known as genetic variance of environmental sensitivity, exists for SCC. The importance of reported G×E for SCS is limited. Sire × herd interaction effects on SCS explained between 0 and 3% of the total phenotypic variance (Banos and Shook, 1990; Schutz et al., 1994; Samore et al., 2001). Estimated genetic correlations between SCS expressed in herd environments with low versus high average SCC were mainly close to unity (Castillo-Juarez et al., 2000; Raffrenato et al., 2003; Calus et al., 2005c), apart from a reported value of 0.80 estimated for

Swedish Holstein (Carlén et al., 2005), and a value as low as 0.83 when environments were defined based on management practices that enhance milk production (Raffrenato et al., 2003). All these studies were based on lactation average SCS, and no results are available on G×E of SCS based on test-day records. The G×E in SCS of animals on test-day level might be stronger due to short term changes in the environment, such as increased incidence of mastitis infection.

In this paper, we investigated the magnitude of genotype, environment, and G×E for SCS, related to herd environment based on herd test-day BMSCC, DIM, and the interaction of BMSCC and DIM of the individual cow.

MATERIALS AND METHODS
Data

In total 6,770,924 test-day records were available from Dutch dairy herds during 1997, 1998, and 1999, including repeated lactations. All animals were at least 75% Holstein Friesian. To reduce the number of records, randomly 25% of all herds were selected, reducing the number of records to 1,663,898. Herds needed at least 20 records on each herd test-day. Records before 5 DIM and after 365 DIM were deleted, as well as records of animals with fewer than 5 test-day records. This last criterion was applied to avoid bias due to inclusion of incomplete lactation records in the analysis (Pool and Meuwissen, 2000). Records deleted in this step had an average SCC of 207,000 cells/mL, whereas the remaining records had an average SCC of 186,000 cells/mL. This indicates that the deleted records, had a higher than average proportion of affected records. Further, records of animals calving for the first time at an age of less than 640 d were deleted, as well as records of parity 5 and higher. These editing steps reduced the number of test-day records to 1,087,635 (28,322 herd test-days). Additional editing steps deleted sires with fewer than 25 daughters, sires with daughters in fewer than 3 herd test-days and herd test-days with daughters of fewer than 3 sires. Finally, herd test-days with fewer than 5 remaining records were deleted. The final data set contained 696,826 test-day records of 49,130 animals in 947 herds on 27,532 herd test-days. For each herd test-day, BMSCC was calculated as average of all available SCC records on that herd test-day, weighted by individual milk production. Somatic cell score was calculated from SCC (SCS = $\log_2(SCC/100,000)+3$).

For estimation of variance components and breeding values using a sire model, the final data set was halved by randomly selecting half of the herds, so that daughter performance in the other 50% of the final data set could be used to check the predictive ability of the sires PTA. This last step selected 344,029 test-day records of 24,125 cows in 461 herds on 13,563 herd test-days (the other half contained 352,797 test-day records). The pedigree included 479 animals of which 182 were sires with daughters in the data.

Random regression model

Variance components were estimated with a sire model, assuming that SCS was the same trait in different lactations apart from the fixed effect corrections. Random effects were included for sire, and two effects for cow. The genetic sire effect was modelled by applying random regressions (**RR**) i) on DIM, to account for differences in lactation curves, and ii) on herd test-day BMSCC, to account for differences in environmental sensitivity to changing BMSCC, and in a second model also iii) on the interaction between BMSCC and DIM to account for specific differences in lactations curves in environments with different BMSCC. The within lactation animal effect was modelled by applying RR (for each lactation separately) on DIM, to account for individual differences in lactation curves, and on herd test-day BMSCC, to account for change in variances with changing BMSCC. The random between lactation animal effect was modelled by random effects for each animal. The RR were applied to Legendre polynomial coefficients (Kirkpatrick et al., 1990) representing BMSCC, DIM, and the interaction between them. Heterogeneous residual variances were included in the model for 25 groups that were formed by first splitting the data in 5 equally sized groups based on increasing BMSCC, and then splitting the data in 5 equally sized groups based on increasing DIM. Residual covariances other than permanent environmental were assumed to be zero. The residual groups contained between 8166 and 17,219 test-day records. To account for within residual group averages, a fixed effect was added for each residual group as well. Other fixed effects were included in the model for mean, year-season, parity, and herd test-day. Fixed regressions were included to account for age at calving within parity, breed of the cow, for DIM within parity, and for the interaction between DIM and BMSCC. No fixed regression on BMSCC was included as effects of BMSCC were accounted for by the fixed effect for herd test-day.

The model was:

$$Y_{iklnpq} = \mu + FIXED\ EFFECTS + \sum_{j=0}^{10} \beta_{ij} P_{ijl} + \sum_{j=0}^{10} \gamma_j Q_{ijklq} +$$

$$\left[\sum_{m=0}^{s} \alpha_{mn} R_{imq} + \sum_{o=1}^{t} \varphi_{on} S_{ioq} + \lambda_n T_{ioq} \right] + \left[\sum_{m=0}^{s} \varpi_{imq} R_{imq} + \sum_{o=1}^{t} \rho_{ioq} S_{ioq} \right] + animal_q + E_{iklnpq}$$

where

Y_{iklnpq} is an SCS record of cow q, μ is the average performance over all animals, *FIXED EFFECTS* included year-season, herd test-day, residual group, and second order polynomial regressions on age at calving and percentage of Holstein Friesian, Dutch Friesian, and Meuse-Rhine-Yssel genes, $\sum_{j=0}^{10} \beta_{ij} P_{ijl}$ is a fixed 10th-order regression within parity i (1,2, …,4) (β_{ij})

on a polynomial coefficient reflecting DIM l (P_{ijl}), resembling the average lactation curve in the population, $\sum_{j=0}^{10} \gamma_j Q_{ijklq}$ is a fixed 10th-order regression (γ_j) on a polynomial coefficient resembling the interaction of BMSCC at herd test-day k and DIM l of cow q (Q_{ijklq}), α_{mn} is coefficient m of the RR on the orthogonal polynomial coefficients of herd test-day BMSCC of the daughters of sire n, φ_{on} is coefficient o of the RR on the orthogonal polynomial coefficients of DIM of the daughters of sire n, λ_n is the coefficient of the linear RR on the orthogonal polynomial coefficients of the interaction of herd test-day BMSCC and DIM of the daughters of sire n, ω_{imq} is coefficient m of the RR on the orthogonal polynomials coefficients of herd test-day BMSCC of cow q in parity i (permanent environment within lactation), ρ_{ioq} is coefficient o of the RR on the orthogonal polynomials coefficients of DIM of cow q in parity i (permanent environment within lactation), s and t are the largest significant estimable coefficients m and o of the RR on BMSCC and DIM, respectively, R_{imq}, S_{ioq} and T_{ioq} are polynomial coefficients reflecting DIM, BMSCC and the interaction between them of cow q in parity i, $animal_q$ is a random effect correcting for between lactation permanent environmental variance of cow q, and E_{iklnpq} is the residual effect of cow q in herd test-day k within residual group p (p = 1, 2, ..., 25).

In matrix notation, the model was: $\mathbf{y} = \mathbf{Xb} + \{\mathbf{As_{DIM}} + \mathbf{Bs_{BMSCC}} + \mathbf{Us_{BMSCC*DIM}}\} + \{\mathbf{Vp_{DIM}} + \mathbf{Wp_{BMSCC}}\} + \mathbf{Yp_{animal}} + \mathbf{Ze_{BMSCC,DIM}}$, where $\mathbf{Xb}$ represents all fixed effects, $\mathbf{As_{DIM}}$, $\mathbf{Bs_{BMSCC}}$, and $\mathbf{Us_{BMSCC*DIM}}$ the additive genetic effects, $\mathbf{Vp_{DIM}}$, and $\mathbf{Wp_{BMSCC}}$ the within lactation permanent environmental effects, $\mathbf{Yp_{animal}}$ the between lactation permanent environmental effects, and $\mathbf{Ze_{BMSCC,DIM}}$ the residual effects. Variances and covariances were modeled across BMSCC and DIM, for additive genetic effects, and permanent environmental effects. Residual variances were estimated for each residual group, and residual covariances between groups were assumed to be zero. Heritabilities were calculated as 4 times the sire variance divided by the sum of the residual variance, the within and between lactation permanent environmental variance and the sire variance. All analyses were performed with ASReml (Gilmour et al., 2002b).

Stepwise increasing of orders of RR

The RR, modelling sire and cow effects, were stepwise increased. At first, first order RR on BMSCC for sire and cow were included in the model. The order of the RR on DIM were increased for sire and cow effects together, until the highest order was not significantly estimable. The model was both applied with and without a RR for sire effects on the interaction between BMSCC and DIM. After reaching the highest order for the RR on DIM, finally the order of the RR on BMSCC was attempted to increase further. Likelihood ratio tests were used to identify the highest estimable significant orders for the sire and cow effects

(P < 0.05). The test statistic was twice the difference in log likelihood between models with order n and n-1, respectively.

Bivariate model

A bivariate model was applied to enable comparison of estimated variances and correlations of the random regression model (**RRM**). In the bivariate model, SCS was considered to be a different trait in each of the defined residual groups in the RRM (based on increasing BMSCC and DIM). If animals had more than one record in a residual group, one was randomly selected and included in the analysis. The bivariate model contained for both traits fixed effects including average performance, year-season, herd test-day, a 10th-order polynomial regression on DIM within each parity and second order polynomial regressions on age at calving and percentage of Holstein Friesian, Dutch Friesian, and Meuse-Rhine-Yssel genes. For both traits residual and sire variances of the traits were estimated, as well as the residual and sire covariances between the traits.

Correlated responses across environments

Correlated responses across environments were investigated for a situation reflecting selection solely on sire PTA for SCS in one environment and the response in another environment. The assumption was that for each combination of BMSCC and DIM, one genetic standard deviation of genetic progress was made. Based on the estimated genetic variances and correlations, correlated responses were calculated across BMSCC and DIM, as $CR_{Dx1,By1} = ir_{(Dx1,By1),(Dx2,By2)}\sigma_{Dx1,By1}$, where $CR_{Dx1,By1}$ is the correlated response in SCS at DIM x_1 and BMSCC y_1, i is the selection intensity (set to one genetic standard deviation for each situation), $r_{(Dx1,By1),(Dx2,By2)}$ is the additive genetic correlation between SCS at DIM x_1 and BMSCC y_1, and SCS at DIM x_2 and BMSCC y_2, and $\sigma_{Dx1,By1}$ is the additive genetic standard deviation of SCS at DIM x_1 and BMSCC y_1 (Falconer and Mackay, 1996).

Phenotype versus BMSCC and PTA

The combined effects of BMSCC and genetic merit of sires on phenotypic SCC were estimated by fitting the RRM (with the highest estimable orders for the RR) with a 10th-order fixed polynomial regression on BMSCC, instead of a fixed herd test-day effect. Average estimated breeding values were zero within environments, and hence the average phenotypic performance estimated with the 10th-order regression on BMSCC was combined with PTA of zero. Within a herd environment, the change in phenotypic performance was calculated as a correlated response to selection in an environment with an average BMSCC of 184,000 cells/mL, reflecting the effects in herds of different BMSCC of selecting sires on a national index for SCS. The considered range of sires PTA was 2 sire standard deviations. The considered PTA were estimated at an average stage of lactation of 167 DIM.

Table 7.1. *Log likelihoods of the fitted models with orders of the random regression (RR) for sire and cow effects on bulk milk somatic cell count (BMSCC) and days in milk (DIM), and without or with a first order RR on their interaction (BMSCC*DIM).*

Order RR		Log likelihood	LRT
BMSCC	DIM		BMSCC*DIM
1	1	-1737143.71	46.66
1	2	-1732505.80	58.20
1	3	-1730494.83	53.73
1	4	-1729550.69	56.37

[1]Loglikelihood ratio test (LRT) statistic calculated as twice the difference in log likelihood of models with and without the linear regression on the interaction between BMSCC and DIM. Differences between the models are significant if LRT statistic > 3.84.

RESULTS
Model selection

The RRM that best fitted the data based on the log likelihood ratio test included a fourth order RR on DIM, and first order RR on BMSCC and on the interaction between BMSCC and DIM (Table 7.1). Models with higher order RR on DIM and BMSCC did not converge. Differences in log likelihood were larger for pairs of models with increasing order on DIM, than models with and without a RR on the interaction between DIM and BMSCC.

The pattern of the sire variance did change between models with and without a RR on the interaction between BMSCC and DIM (Figure 7.1). The largest differences were found in estimates early in lactation on herd test-days with a high BMSCC, where the estimated sire variances for the model with the RR on the interaction between BMSCC and DIM were nearly twice as high as the estimates of the model without that effect. Estimated sire variances of the bivariate model showed generally the same trends across DIM and BMSCC as the RRM (results not shown). All reported results are from the RRM including a fourth order RR on DIM, and first order RR on BMSCC and on the interaction between BMSCC and DIM, unless stated otherwise.

Genetic parameters for SCS across BMSCC and DIM

Estimated heritabilities for SCS were highest early in lactation at high BMSCC and late in lactation at low BMSCC (Table 7.2). Heritabilities for SCS were lowest early in lactation at low BMSCC (Table 7.2). Estimated genetic correlations between SCS measures at extreme DIM in the same environments ranged from 0.43 to 0.65 (Table 7.3). Trends in estimated genetic correlations were comparable between pairs of DIM across BMSCC, but correlations dropped in most cases where the difference in BMSCC increased (Table 7.3). Genetic correlations between SCS in environments with extreme BMSCC became as low as 0.72 early
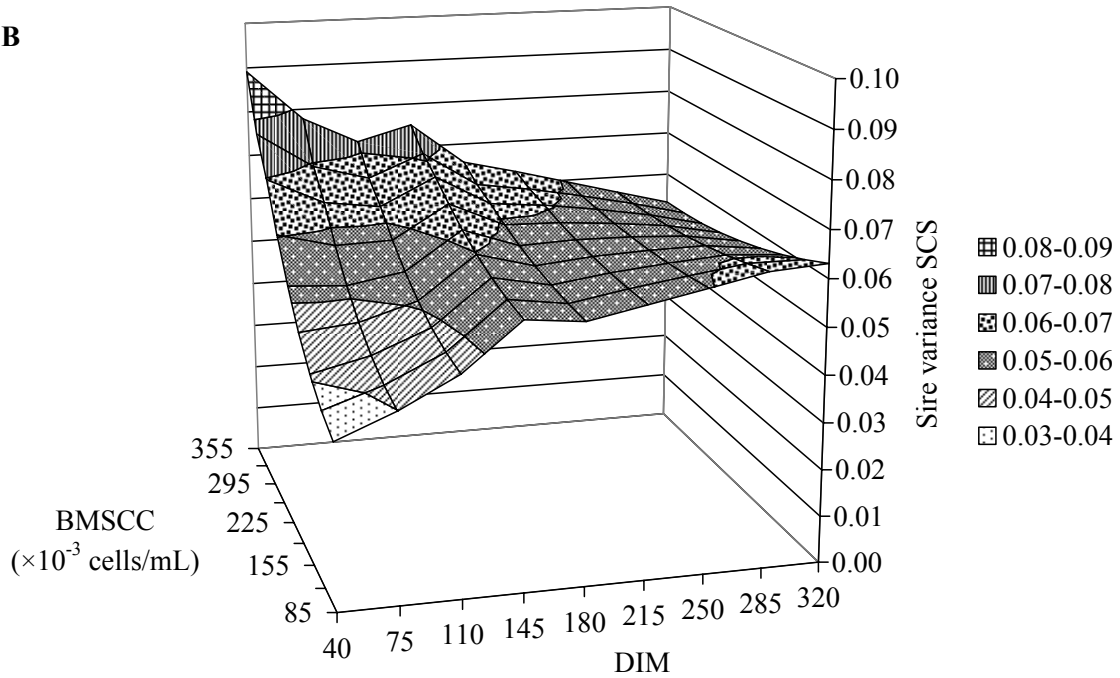
A
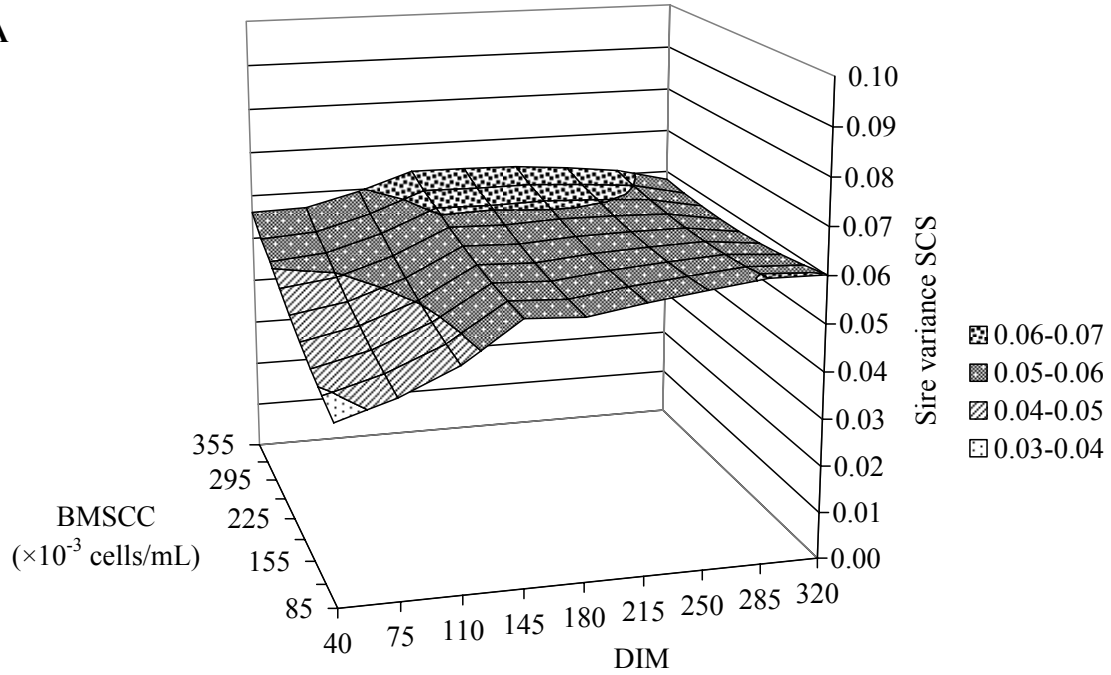
B

Figure 7.1. *Sire variances of SCS on herd test-days with different bulk milk somatic cell count (BMSCC) and at different days in milk (DIM), estimated with a random regression model (RRM) with a fourth order random regression on DIM and first order RR on BMSCC (**A**) or with a fourth order RR on DIM and first order RR on BMSCC and the interaction between BMSCC and DIM (**B**).*

in lactation, but were close to unity late in lactation. The lowest estimated genetic correlation was 0.42, between a situation early in lactation at high BMSCC and a situation late in lactation at low BMSCC. Most genetic correlations between SCS at different DIM and BMSCC were comparable between the RRM and bivariate model in situations where one of both environments had a BMSCC of 175,000 cells/mL or less (Table 7.3). However, estimated correlations of the bivariate model for situations where both environments had a BMSCC of 230,000 or 360,000 cells/mL tended to be closer to unity.

Table 7.2. *Estimated heritabilities[1] of SCS on herd test-days with different bulk milk somatic cell count ($\times 10^{-3}$ cells/mL) (BMSCC) and at different days in milk (DIM), estimated with a random regression model with a fourth order random regression (RR) on DIM and first order RR on BMSCC and the interaction between BMSCC and DIM.*

BMSCC	DIM				
	40	110	175	245	315
85	0.06	0.08	0.11	0.14	0.16
130	0.06	0.08	0.11	0.13	0.15
175	0.06	0.08	0.10	0.13	0.14
230	0.07	0.08	0.10	0.11	0.13
360	0.10	0.09	0.10	0.10	0.11

[1]Approximate standard errors of the heritabilities ranged from 0.01 to 0.03.

Sires PTA for SCS across DIM and BMSCC

To gain insight in the differences in patterns of PTA of sires across BMSCC, the PTA for SCS of the ten sires with most daughter records in the data were plotted across BMSCC at 40 and 315 DIM (Figure 7.2). Both at 40 and 315 DIM, the response of the sires PTA to increasing BMSCC was that some of the sires PTA decreased while some of the sires PTA increased. Comparing sire A (marked with triangles) and sire B (marked with squares), sire B had a more desirable PTA across BMSCC (i.e., a lower value) early in lactation, while sire A had a more desirable PTA across BMSCC late in lactation. The breeding value of sire A decreased with increasing BMSCC early in lactation, while it increased with increasing BMSCC late in lactation.

Selection response for SCS across DIM and BMSCC

Correlated selection responses in all response environments were calculated based on selection in all environments (Table 7.4). Values for BMSCC and DIM in each row indicate the selection environment, while values for BMSCC and DIM in each column indicate the environment of the selection response. The diagonal represents the direct response to selection in the selection environment, which was arbitrarily chosen to be one sire standard

Table 7.3. *Estimated genetic correlations[1] between SCS on herd test-days with different bulk milk somatic cell count ($\times 10^{-3}$ cells/mL) (BMSCC) and at different days in milk (DIM), estimated with a random regression model (RRM) with a fourth order random regression on DIM and first order random regression on BMSCC and the interaction of BMSCC and DIM or a bivariate model.*

RRM

	bmscc	85	85	130	130	130	175	175	175	230	230	230	360	360	360
bmscc	dim	175	315	40	175	315	40	175	315	40	175	315	40	175	315
85	40	0.79	0.56	0.98	0.80	0.58	0.95	0.81	0.60	0.88	0.81	0.62	0.72	0.80	0.65
85	175		0.87	0.76	1.00	0.89	0.71	0.99	0.90	0.65	0.97	0.92	0.51	0.91	0.93
85	315			0.55	0.86	1.00	0.53	0.85	0.99	0.50	0.83	0.98	0.42	0.76	0.92
130	40				0.78	0.57	0.99	0.80	0.59	0.95	0.82	0.60	0.83	0.83	0.63
130	175					0.88	0.75	1.00	0.90	0.69	0.99	0.91	0.56	0.94	0.93
130	315						0.55	0.87	1.00	0.51	0.85	0.99	0.42	0.78	0.95
175	40							0.78	0.56	0.99	0.80	0.58	0.91	0.85	0.60
175	175								0.89	0.73	1.00	0.90	0.61	0.96	0.92
175	315									0.52	0.87	1.00	0.43	0.80	0.96
230	40										0.77	0.53	0.96	0.84	0.54
230	175											0.88	0.67	0.98	0.91
230	315												0.43	0.82	0.98
360	40													0.77	0.43
360	175														0.85

Bivariate

	bmscc	85	85	130	130	130	175	175	175	230	230	230	360	360	360
bmscc	dim	175	315	40	175	315	40	175	315	40	175	315	40	175	315
85	40	0.96	0.84	0.84	0.83	0.81	0.79	0.87	0.81	0.65	0.79	0.84	0.52	0.74	0.51
85	175		0.82	0.58	0.91	0.71	0.42	0.69	0.53	0.45	0.62	0.58	0.23	0.53	0.48
85	315			0.68	0.86	0.93	0.52	0.74	0.67	0.46	0.44	0.55	0.01	0.43	0.44
130	40				0.80	0.86	0.96	0.97	*1.00*	0.92	0.93	1.00	*1.00*	0.96	0.96
130	175					*1.00*	0.67	0.97	*1.00*	0.64	0.85	0.71	0.29	0.64	0.71
130	315						0.68	0.95	0.99	0.67	0.80	0.80	0.45	0.60	0.81
175	40							0.97	0.95	*1.00*	0.93	*1.00*	0.99	1.00	0.99
175	175								0.99	*1.00*	*1.00*	0.94	*1.00*	*1.00*	*1.00*
175	315									0.91	0.86	0.88	*1.00*	0.85	0.99
230	40										0.98	0.92	*1.00*	*1.00*	*1.00*
230	175											0.76	*1.00*	0.97	0.99
230	315												0.75	0.59	0.93
360	40													*1.00*	*1.00*
360	175														*1.00*

[1]Approximate standard errors ranged from 0.01 to 0.13 for the RRM and from 0.04 to 0.32 for the bivariate model. Estimated genetic correlations printed in italic for the bivariate model were fixed at the boundary.
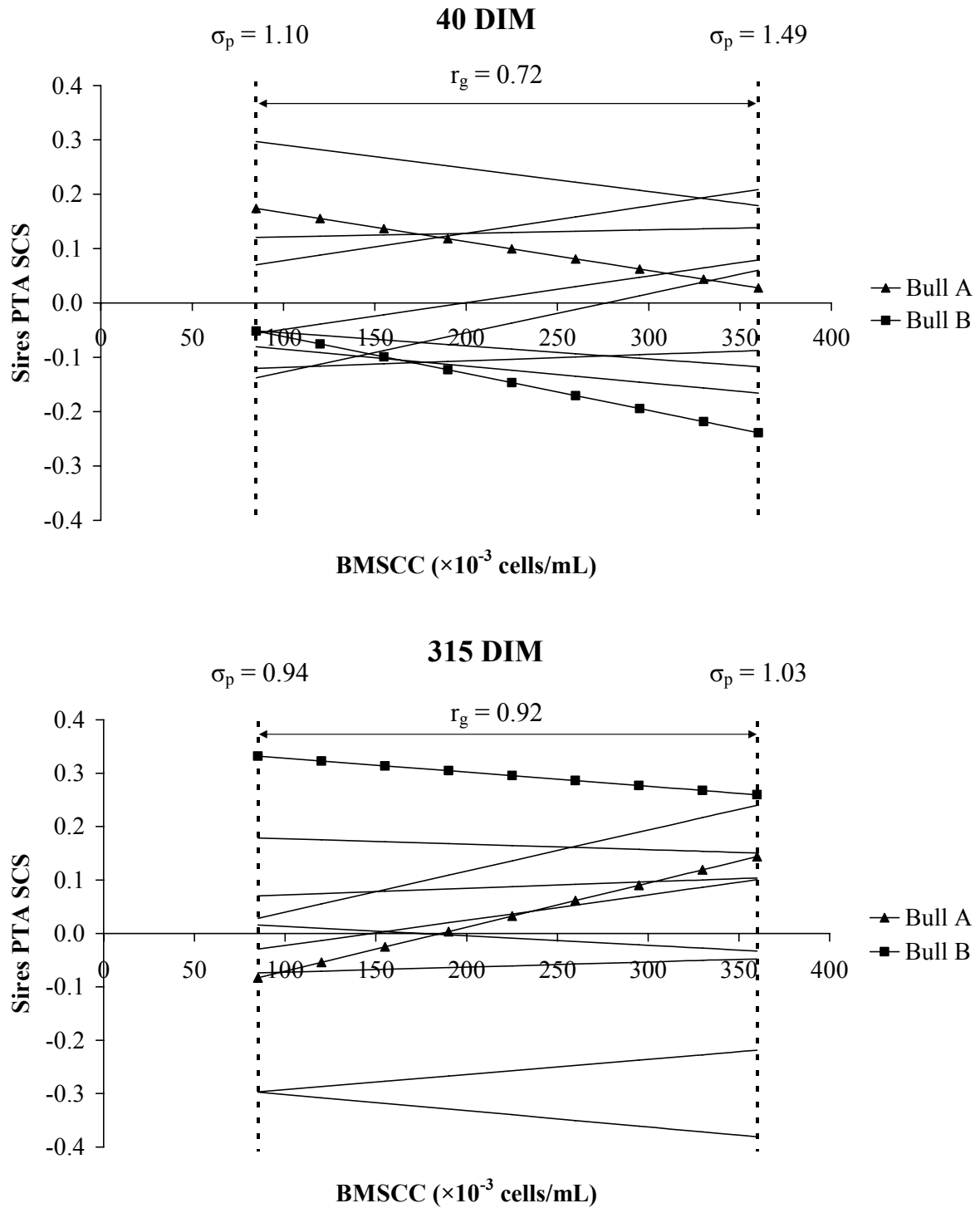
Figure 7.2. *Predicted transmitting abilities estimated for sires with most daughter records in the data, across BMSCC at DIM of 40 and 315 d. Phenotypic standard deviations are given for the tenth and 90th percentiles of the data (indicated by dotted lines) together with the genetic correlation of SCS between them.*

Table 7.4. *Calculated correlated responses of SCS at different values of bulk milk somatic cell count ($\times 10^{-3}$ cells/mL) (BMSCC) and at different days in milk (DIM). The environment of the correlated response is characterised by column values for BMSCC and DIM, and the selection environment is characterised by row values for BMSCC and DIM. Responses to selection in the selection environment are given on the diagonal.*

Selection environment		Response environment								
	BMSCC	85	85	85	175	175	175	360	360	360
BMSCC	DIM	40	175	315	40	175	315	40	175	315
85	40	0.187	0.184	0.140	0.197	0.191	0.144	0.219	0.205	0.151
85	175	0.147	0.233	0.219	0.149	0.234	0.217	0.152	0.234	0.214
85	315	0.104	0.203	0.251	0.111	0.201	0.239	0.126	0.196	0.214
175	40	0.177	0.166	0.134	0.209	0.183	0.135	0.274	0.218	0.138
175	175	0.151	0.231	0.214	0.162	0.236	0.213	0.185	0.248	0.213
175	315	0.111	0.211	0.250	0.117	0.209	0.241	0.129	0.207	0.223
360	40	0.135	0.118	0.105	0.190	0.145	0.103	0.302	0.200	0.099
360	175	0.149	0.212	0.191	0.176	0.227	0.193	0.233	0.258	0.197
360	315	0.122	0.216	0.233	0.124	0.217	0.232	0.129	0.219	0.232

deviation. The lowest correlated responses were less than half as large as the direct response to selection in the environment itself. For selection early in lactation (40 DIM), correlated responses were lowest late in lactation (315 DIM). For selection later in lactation (175 and 315 DIM), the lowest correlated responses were estimated in an environment with BMSCC of 85,000 cells/mL and 40 DIM. Correlated responses across BMSCC tended to be lowest in herds with low BMSCC and highest in herds with high BMSCC. The correlated responses did show similar trends as the genetic correlations (Table 7.3). However, the differences in genetic variance across DIM and BMSCC also had an influence on the correlated responses, as selection in an environment with low BMSCC generally yielded a higher correlated response in environments with high BMSCC, rather than vice versa.

DISCUSSION

Breeding values were estimated depending on DIM and BMSCC. As a result, estimated breeding values represented a surface across DIM and BMSCC, rather than one single point when DIM and BMSCC are ignored, or a line when for instance only a RR on DIM is performed. The RR on DIM was chosen to account for individual differences in SCS patterns across DIM. The RR on BMSCC was chosen, since BMSCC is related to the hygienic

conditions on the farm (Barkema et al., 1999b). Including two dimensions in the model, i.e., a fourth order RR on DIM and a linear RR on BMSCC, implied that an animals pattern of the EBV for SCS across DIM was linearly scaled across BMSCC. A second model that also included a RR on the interaction between BMSCC and DIM, did allow individual SCS patterns across DIM to be different across BMSCC. The results indicated that across DIM and herd environment considerable G×E exists for SCS. This G×E might come from different sources of variation, e.g., type of mastitis and incidence of mastitis, being involved in the genetic variation of SCC, and will have implications for the optimal selection strategies in breeding programs. Also, at farm level this G×E has implications in that management and genetic selection need to be considered together and not as separate components for the reduction of SCC. Different sources of variation for SCC, the implication of G×E for breeding programs, and the optimal balance between management and genetic selection for the reduction of SCC are discussed below.

Genetic variation for G×E on SCS across DIM and BMSCC

The increase of the estimated heritability of SCS across DIM was in line with results of other studies applying test-day models to Holstein cow data (Haile-Mariam et al., 2001; Odegard et al., 2003; Koivula et al., 2004). Two of these studies found a similar increase in the genetic variance across DIM (Odegard et al., 2003; Koivula et al., 2004), although others found a stronger increase (Haile-Mariam et al., 2001). Estimated genetic correlations between extreme DIM ranged from 0.43 to 0.63, which was in line with reported values ranging from 0.3 to 0.7 (Haile-Mariam et al., 2001; Odegard et al., 2003; Koivula et al., 2004).

The defence of cows against mastitis pathogens is reported to be lower early compared with late in lactation (Mallard et al., 1998). The lower defence early in lactation likely results in a higher incidence of mastitis, which might be one of the reasons why there was higher genetic variance for SCS early in lactation. However, a threefold higher genetic variance was found also early in lactation in environments with high BMSCC compared with environments with low BMSCC. Low BMSCC is associated with a higher risk of mastitis due to environmental pathogens such as *Escherichia coli* (Erskine et al., 1988; Miltenburg et al., 1996; Barkema et al., 1998b). Thus, the low estimated sire variance early in lactation at low BMSCC might indicate that there was little difference in susceptibility to environmental pathogens early in lactation between daughters of different sires. High BMSCC is associated with a higher risk of mastitis due to contagious pathogens such as *Streptococcus agalactiae* and *Staphylococcus aureus* (Erskine et al., 1988; Wilson et al., 1997; Barkema et al., 1998b). Therefore, the high estimated sire variance early in lactation at high BMSCC might be caused by some animals being infected and some animals not infected, indicating that there are large genetic differences in susceptibility to contagious pathogens early in lactation between daughters of different sires. Contagious pathogens lead to elevated SCC for several test-days

(De Haas et al., 2002). In herds with high BMSCC and contagious pathogens a relatively large part of the animals might have had an elevated SCC towards the end of the lactations. This might explain that differences between animals (i.e. genetic variance) were lower towards the end of the lactation in high BMSCC herds. Hence, in early lactation the difference in susceptibility is the major cause of variation, and at the end of lactation the most important source of genetic variation are differences in SCS among affected animals. The estimated genetic correlation early in lactation between SCS at low and high BMSCC (i.e., 0.72) support that different sources of variation might be active in herds with low and high BMSCC.

Genetic correlations of SCS across environments, estimated on a lactation base, are reported to be between 0.8 and unity (Castillo-Juarez et al., 2000; Raffrenato et al., 2003; Calus et al., 2005c; Carlén et al., 2005). Our genetic correlations, estimated on a test-day base, indicated more reranking of sires across BMSCC early in lactation and comparable or less reranking of sires late in lactation (i.e., correlations ranged from 0.72 to 1.00). This further supports the idea that G×E for SCS is more important early in lactation, possibly due to the higher incidence of mastitis. It also shows that a more detailed analysis of phenotypic information, e.g., on test-day basis rather than lactation averages, reveals more G×E. In this specific case, the use of herd test-day specific BMSCC would rather reflect temporal environmental changes, such as outbreaks of mastitis, whereas a herd-year average BMSCC probably would rather reflect average herd management.

Implications for breeding programs

Based on the estimated G×E for SCS between environments with different BMSCC and at different DIM, it can be argued that the breeding goal to reduce SCS should depend on BMSCC and DIM. Another strategy might be to have one breeding goal for all environments and emphasize selection for reduced SCS in those circumstances where reducing SCS is most important. Identification of those circumstances brings us back to the main aims to reduce SCS through selection: 1) reduce incidence of mastitis by using SCS as a predictor trait (Emanuelson, 1988; Weller et al., 1992; Philipsson et al., 1995), and 2) decrease the chance to get a penalty for high BMSCC (Dekkers et al., 1996; Veerkamp et al., 1998). The incidence of mastitis is usually highest early in lactation (Erskine et al., 1988; Barkema et al., 1998b), and is also strongly correlated to SCS early rather than late in lactation (De Haas et al., 2003). Thus, early in lactation, elevated SCS might be especially important as an indicator for incidence of mastitis. Reducing SCS as a way to decrease BMSCC is likely more important at high BMSCC, as the chance to get a penalty is higher. Therefore, the major focus in a breeding goal could be to decrease SCS early in lactation and at high BMSCC. Following this strategy, it should be taken into consideration that the selection using breeding values for the average environment (BMSCC at 175 and DIM at 175) gives a response in early lactation at

high BMSCC, as low as 60% of the possible selection response (Table 7.4). The results in Table 7.4 indicate that for selection for reduced SCS at high BMSCC early in lactation, emphasis should be on environments early in lactation and in environments with average and above average BMSCC. In these circumstances, heritabilities ranged from 0.06 to 0.10 (medium to high BMSCC). The reliability of a sires EBV is $r^2 = n_e/(n_e + \lambda)$, where n_e is the effective daughter size and $\lambda = (4-h^2)/h^2$. In a situation where selection is on one trait, while the response is on another trait, h^2 is replaced by the coheritability, which is calculated as $h_X h_Y r_A$, where h_X is heritability of trait X, h_Y is heritability of trait Y, and r_A is the genetic correlation between both traits. This implies that in order to get a breeding value with a reliability of 80% for SCS early in lactation at medium or high BMSCC, respectively 251 or 255 daughters are required in an average environment (BMSCC at 175 and DIM at 175). From the same formulas, it follows that with direct selection early in lactation at medium or high BMSCC, respectively 251 and 157 daughter records are needed to get sires breeding values for those circumstances with a reliability of 80%. This illustrates that evaluating sires based on an average environment reduces accuracy in those circumstances where decreasing SCS might be most important.

Reducing SCS at farm level

Both management and genetic selection can be used to reduce SCS at farm level, and the existence of G×E indicates that both should be considered simultaneously in order to evaluate their relative importance. To enable this, the average phenotypic performance for SCS was estimated as a function of BMSCC and sires PTA for SCS (Figure 7.3). The PTAs for sires were based on the average environment, but the effects of selection were calculated for each environment specifically, thus depended on the genetic correlation and genetic SD in each BMSCC environment. Figure 7.3 shows that both at high and low level of BMSCC there are considerable benefits of using the best sire for SCS. Sometimes it is argued that at high levels of BMSCC farmers should first take management action before considering breeding for reduced SCC. Although this might be the quickest solution in the short term, Figure 7.3 shows that at high BMSCC the benefits of selecting the sires with best PTA for SCC are considerable, and as discussed above even higher than when selecting at average or low BMSCC.

Figure 7.3 also shows that for different values of BMSCC, a decrease in BMSCC does not lead to comparable changes in average SCS. This is partly a consequence of the different scales of BMSCC and SCS (being log-transformed or not), and of the fact that BMSCC is an average of SCC weighted by daily milk production of the animals while the average phenotypic SCS is not weighted. However, BMSCC (which is not log-transformed) is preferred as it is a widely known measure for the environment, while SCS (which is log-transformed) is preferred because of its statistical properties. The difference in scale could be

solved by calculating PTA directly on the scale of SCC, ignoring non-normality of SCC. Comparison of PTA on the scale of SCC and SCS, both depending on BMSCC and DIM, indicated that PTA on the scale of SCC showed larger G×E effects and actually better predicted average daughter performance, than did PTA on the scale of SCS (Calus et al., 2005b). The log-transformation might result in loosing some important information, since the differences between records with high SCS records (i.e., records that are likely affected by mastitis) are mainly affected. A better solution might be to consider SCS records of animals that are either affected or not with mastitis, as different traits. Application of a reaction norm model could partly solve this problem, where the difference of affected and non-affected records can be explained as a G×E effect, but other solutions, such as application of mixture models (Detilleux and Leroy, 2000) have been proposed.
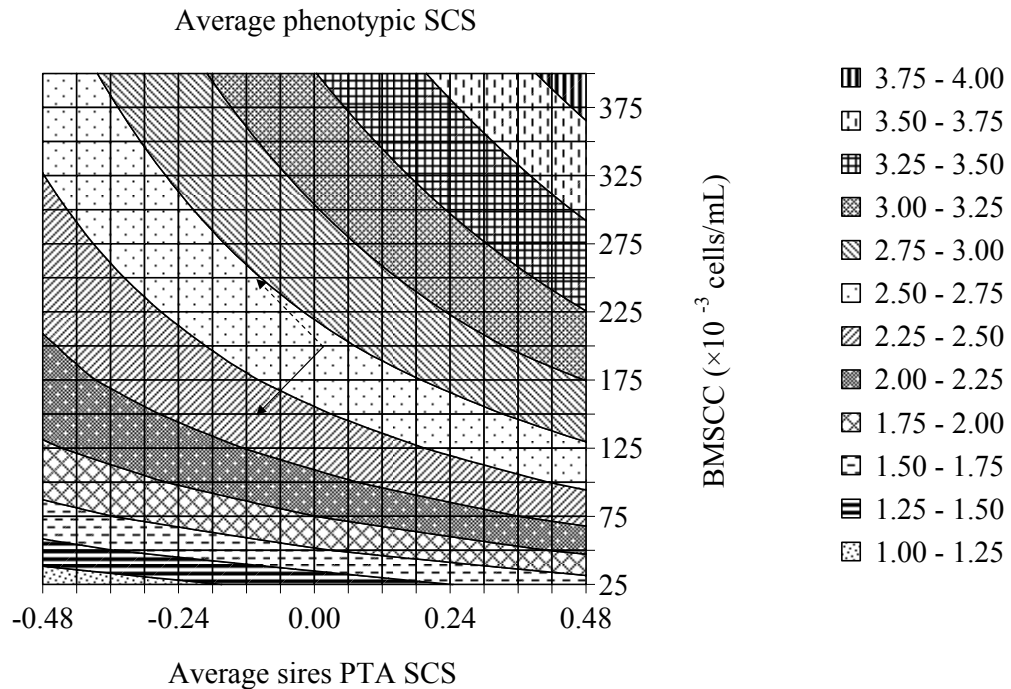
Average phenotypic SCS

Figure 7.3. *Average phenotypic performance for SCS, plotted as function of BMSCC and average sires PTA for SCS, assuming that sires are selected based on their breeding values in an environment with average BMSCC (184,000 cells/mL) and average stage of lactation (167 DIM). The arrows indicate changes in SCS in situations with selection for reduced SCS combined with decreasing (solid arrow) and increasing BMSCC (dotted arrow).*

CONCLUSIONS

Greater G×E effects were estimated on a test-day basis, than reported G×E effects based on lactation averages, indicating that more detailed analysis of phenotypic information reveals more G×E. Early in lactation a strong G×E effect was detected for SCS: between herds with low and high BMSCC the genetic variance increased threefold and reranking of sires occurred. Early in lactation, heritabilities were highest at high BMSCC, indicating more accurate testing of bulls under these circumstances. Responses to selection were reasonably high among environments in the second half of the lactation, while responses to selection between environments early and late in lactation tended to be low. Selection for reduced SCS yielded the highest direct response early in lactation at high BMSCC.

ACKNOWLEDGEMENTS

Chapter 8

Genetic correlations between milk production and health and fertility dependent on herd environment

J. J. Windig

M. P. L. Calus

R. F. Veerkamp

Animal Sciences Group

P.O. Box 65, 8200 AB Lelystad, The Netherlands

ABSTRACT

High milk production in dairy cattle can have negative side effects on health and fertility traits. This paper explores the genetic relationship of milk yield with health and fertility dependent on herd environment. A total of 71,720 lactations from heifers calving in 1997 to 1999 in the Netherlands were analyzed. Herd environment was described by 4 principal components: intensity, average fertility, farm size and relative performance indicating whether herds had good (poor) health and fertility despite a high (low) production. Fertility was evaluated by days to first service (DFS) and number of inseminations (NINS) while somatic cell score (SCS) was used as a measure of udder health. Data were analyzed with a multi-trait reaction norm model. Genetic correlation within traits across environments ranged from 0.84 to unity. Genetic correlations of the three traits with milk yield were antagonistic but varied over environments. Genetic correlation of milk yield with DFS varied from 0.30 in small herds to 0.48 in herds with low average fertility. Correlations with NINS varied from 0.18 in large herds to 0.64 in high fertility herds, and with SCS from 0.25 in herds with a high fertility relative to production to 0.47 in herds with a relatively low fertility. Selection in environments of average value resulted in different predicted responses over environments. For example, selection for a decrease of NINS of 0.1 in an average production environment decreased milk yield by 35 kg in low but by 178 kg in high production herds.

INTRODUCTION

Breeding for increased production in dairy cattle has negative side effects on health and fertility traits (Pryce et al., 1997; Rauw et al., 1998; Sandoe et al., 1999; Roxstrom et al., 2001). To counter these effects both new breeding goals and management tools have been advised (Neave et al., 1969; Barkema et al., 1999a; Esslemont, 2003; Pryce et al., 2004; Royal and Flint, 2004). Generally management and genetic effects are considered separately. However, genetic parameters, such as genetic correlations between production and health, may change over environments. In other words, selection for an increase in production under one management system may lead to more health risks than under other management systems. Thus management and genetics have to be integrated in order to develop an effective program for improvement of health and fertility.

Although selection for a higher production may on average, lead to more health and fertility problems, there can be considerable variation across herd environments. The phenotypic effect of herd environment (management and other environmental effects) on health and fertility and their relation with milk production at a national scale was recently explored (Windig et al., 2005b; Windig et al., 2005c). With increasing average production days to first service (DFS) and somatic cell score (SCS) decreased, while the number of inseminations (NINS) increased. Within herds, increased production always led to lower fertility and higher SCS. The extent, however, depended on the herd environment. In herds

with a low average production and/or a low average fertility differences between high and low producing animals were relatively small.

In animal breeding, management and other environmental effects are generally accounted for by treating them, as a fixed effect, often in the form of herd-year-season effects. This adjusts results to the average environment, but ignores the interaction between management, environmental effects and the genetics of animals, or the effects on the genetic association between milk yield and health and fertility traits. Recently, reaction norm models which use random regressions to estimate genetic parameters for each environment separately have been explored (Calus et al., 2002; Calus and Veerkamp, 2003; Hayes et al., 2003). A further extension is to model the relationship between several traits across herd environments (Kolmodin et al., 2002; Oseni et al., 2004; Pollott and Greeff, 2004). By doing so genetic correlations between production and health and fertility traits in different herd environments can be estimated.

Objective of this study was to analyze genetic relationships between milk yield and health and fertility traits across herd environments with a multi trait reaction norm model. The overall objective was to assess whether the risks of high milk production in relation to SCS and fertility depended on the herd environment, and what the effect of selection in a specific environment was on traits in the same environment and on traits in other environments.

MATERIALS & METHODS
Data

Test day and insemination data. Data were available for 147,835 first lactation heifers calving between July 1997 and June 1999. Insemination, production and SCC records were the same as in Calus et al. (2005c) and editing steps followed their procedure. All animals were at least 75% Holstein Friesian. Animals were selected if they calved on an age between 640 and 1095 d. Editing steps for environmental parameters (see below) reduced the number of heifers to 116,727. All (grand)daughters of (grand)sires with less than 20 (grand)daughters in the edited data were deleted, reducing the number of animals to 87,375. Herd-year-season subgroups were formed based on the method of Crump et al. (1997) with a minimum of 5 animals per subclass, a minimum length of 30 d and a maximum length of 365 d. Records of animals that could not be assigned to a group with at least five records or that were assigned to a group with less than three records for any of the traits were deleted. Additionally, (maternal grand)sires with progeny in less than three herd-year-season classes and herd-year-season classes with progeny of less than three (maternal grand)sires were deleted. These steps reduced the number of records to 71,270.

Herd environment. The herds were described by 65 environmental variables, partly derived from production data (e.g. average kg fat per cow) and partly from the annual national agricultural survey (e.g. area of the farm). These environmental variables were reduced with

Principal components analysis to four Principal Components (PC's) by choosing the PC's that explained more than 5% of the total variance in all traits (details in Windig et al. (2005c)). The first principal component (PC1, explaining 10.34% of the total variance) was interpreted as production intensity. Environmental variables that contributed most to PC1 were herd averages of 305 d protein, milk and fat yield. Some fertility indices which were negatively correlated to production also contributed negatively to this PC. The second PC (PC2, 8.86%) was interpreted as fertility because all fertility herd averages had large contributions. PC3 (7.20%) combined variables related to the size of the farm such as number of employees, total hours of work, number of animals and area of land. PC4 (6.58%) was a combination of production and fertility, where both production and fertility herd averages had positive contributions. Consequently farms that had a high fertility despite high production received high scores, while farms that performed low both on production and fertility received low scores. Further on PC4 is referred to as relative performance.

Traits analyzed. For production, 305 d milk yield was used. SCC was converted to somatic cell score (SCS) by $SCS = \log_2(SCC/100,000)+3$ (Ali and Shook, 1980) and was used as a health trait. It was defined as the average somatic cell score across test days. Two fertility traits were considered: days to first service (DFS) and number of inseminations per service period (NINS). DFS was calculated as interval from calving to first service. Records for DFS were missing if DFS was smaller than 20 or greater than 300. NINS was missing if NINS was 0 or greater than 10. These editing steps were applied to exclude extremely long lactation records, records with extreme short gestations due to abortions and records with errors.

Pedigree. All sires, paternal grand dams and maternal grandsires of animals with records in the data were included in the pedigree file. All male predecessors of those animals, available from the pedigree data, were also included. Identification of dams of bulls was included if a dam had 2 or more sons and otherwise dams were included as base parents. A total of 1754 animals were included in the relationship matrix.

Reaction norm model

Univariate analysis. Variance components were estimated with a sire-maternal grandsire model. Fixed effects were included in the model for mean and herd-year-season subclass. Fixed regressions were included to account for age at calving and for breed of the cow. The influence of environmental parameters on additive genetic merit was modeled by applying a random regression for each (maternal grand)sire, representing its EBV, on values of a PC for the herd-years in which his (grand)daughters were producing. The incidence matrix of maternal grandsire effects was laid over the matrix of sire effects, i.e. if a bull had both entries in the data as sire and maternal grandsire, the breeding value when being a maternal grandsire was equal to half the breeding value when being a sire. To take heterogeneous residual

variances into account the residual variance was estimated separately for groups of 9000 animals with similar herd environments.

The applied model was:

$$Y_{ijklm} = \mu + FIXED + \sum_{n=0}^{p} \beta_{n,i} \Phi_n(PC_k) + 0.5 \left(\sum_{n=0}^{p} \beta_{n,j} \Phi_n(PC_k) \right) + \varepsilon_{ijklm} \qquad (1)$$

where Y_{ijklm} is the performance of animal *m*, with sire *i*, grand sire *j*, and herd environment *k*, μ is the average performance over all animals, *FIXED* includes herd-year-season subclasses and second order polynomial regressions on age at calving and percentage of Holstein Friesian, Dutch Friesian and Meuse-Rhine-Yssel genes, $\beta_{n,i}$ is coefficient n of the random regression on the orthogonal polynomials of Principal Component Scores of the daughters of sire *i*, $\Phi_n(PC_k)$ (n=0 to p) are the design values of the orthogonal Legendre polynomials of order p for the principal component in environment *k*, $\beta_{n,j}$ is coefficient n of the random regression on the orthogonal polynomials of Principal Component Scores of the maternal granddaughters of sire *j*, ε_{ijklm} is the residual effect of cow *m* in environment *k* within group of environments *l*.

The Legendre polynomials were restricted to a first order polynomial, because of convergence problems for higher order polynomials in the multivariate case, and because higher order polynomials did not give a significantly better fit (see Calus et al. (2005c)). This definition of the genetic model resulted in estimated sire variances of intercepts and slopes and covariances between intercepts and slopes which model possible interactions between slopes and intercepts. From these (co)variances sire variances within single environments could be calculated (Kirkpatrick and Heckman, 1989; Van Tienderen and Koelewijn, 1994; Kolmodin et al., 2002; Oseni et al., 2004).

Residual variances could not be calculated within single environments (i.e. herds) because they contained too few observations (generally in the range of 10 to 20). We calculated residual variances instead for groups of 9000 individuals with similar herd environments. Groups were composed by ranking animals according to their PC-values. The first group consisted of animals 1 to 9000, the second of animal 4501 to 13500, the 3rd of animal 9001 to 18000 etc., till the last group of animals 63001 to 71270 (i.e. the last group contained 8270 animals instead of 9000), resulting in 15 overlapping groups. In order to achieve this grouping the model was run twice, once with animals 1 to 9000, 9001 to 18000, 18001 to 27000, etc. grouped, and once with animals 4501 to 13500, 13501 to 22500, etc. grouped. The additive genetic (co)variances estimated in both runs were very similar and averaged to obtain final esitimates. Heritabilies and genetic correlations (see below under multitrait model) were estimated for the average PC-value of each group (further on pc$_{env1}$ to pc$_{env15}$). Heritabilies were calculated as 4 times the sire variance in pc$_x$ divided by the sum of the residual variance

of the corresponding herd environment group plus 1.25 times the sire variance. The factor 1.25 is explained by the fact that both effects for sires (1 times the sire variance) and maternal grand sires (0.25 times the sire variance) were included in the model.

In some cases the model with intercept and slope did not converge. In that case a reduced model with intercept only was used, and changes in heritabilities were the result of changes in residual variance components only. All analyses were performed with ASREML (Gilmour et al., 2002a).

Multitrait analysis. In order to obtain genetic correlations of production with the other traits in different environments the univariate random regression model was extended to a multitrait analysis. The estimated covariance matrix (V) combined variances and covariances of the random regression, for example, milk and DFS:

$$V = \begin{bmatrix} \sigma^2_{\beta_0(milk)} & & & \\ \sigma_{\beta_0(milk),\beta_1(milk)} & \sigma^2_{\beta_1(milk)} & & \\ \sigma_{\beta_0(milk),\beta_0(DFS)} & \sigma_{\beta_1(milk),\beta_0(DFS)} & \sigma^2_{\beta_0(DFS)} & \\ \sigma_{\beta_0(milk),\beta_1(DFS)} & \sigma_{\beta_1(milk),\beta_1(DFS)} & \sigma_{\beta_0(DFS),\beta_1(DFS)} & \sigma^2_{\beta_1(DFS)} \end{bmatrix} \quad (2)$$

With $\beta_\#$ being the random regression coefficients as defined in equation (1). In order to obtain genetic correlations between the two traits in herd environments pc_1 to pc_{15}, the variance-covariance matrix was computed as M V M' where

$$M = \begin{bmatrix} M_\Phi & 0 \\ 0 & M_\Phi \end{bmatrix} \quad and \quad M_\Phi = \begin{bmatrix} \Phi_0(pc_{env1}) & \Phi_1(pc_{env1}) \\ \Phi_0(pc_{env2}) & \Phi_1(pc_{env2}) \\ \vdots & \vdots \\ \Phi_0(pc_{env15}) & \Phi_1(pc_{env15}) \end{bmatrix} \quad (3)$$

where **0** is a 15 by 2 matrix of zeros and $\Phi_0(pc_{env\#})$ and $\Phi_1(pc_{env\#})$ are the design values of the first order orthogonal Legendre polynomial for herd environment #. In case the multitrait analysis had difficulty with convergence the variances of the traits and the covariance within the traits were fixed to the values of the univariate analyses. In that case only the between trait covariances (i.e. only the four covariances in the lower left hand corner of V) could vary.

Selection response

The implications of the dependency of estimated genetic parameters on the environment were analyzed by calculating the correlated response in SCS and the two fertility traits when selection for an increase in milk took place in pc_8, which is the herd environment at or very close to the average of all herd environments. The correlated response is given by:

$$CR_Y = R_X r_A \frac{\sigma_{A,y}}{\sigma_{A,x}}$$

where CR_Y = the correlated response in trait Y, R_X is the direct response in trait x to selection, r_A is the additive genetic correlation between trait x and y and σ_A is the genetic additive standard deviation (Falconer and Mackay, 1996). The correlated response was calculated for SCS and the fertility traits in pc_{env1}, pc_{env8} and pc_{env15}. So for pc_{env8} the response was in the same environment, while for pc_{env1} and pc_{env15} the response was in another environment and genetic correlation between traits and across environments were used. Correlated responses were also calculated for milk, when selection was directly for SCS, DFS or NINS in pc_8. The response for the directly selected trait was set to +1000kg milk, -6 days for DFS, -0.1 for NINS and -0.5 for SCS. These responses were chosen so that the selection intensity i, is about 2 phenotypic standard deviations, corresponding to approximately 5% of the animals selected in mass selection.

RESULTS
Phenotypic changes

Phenotypically traits varied over the four environmental axes (Principal Components), average production intensity, fertility, farm scale and relative performance (Table 8.1). The range covered by the PC's is given in Table 8.1 for milk production and the health and fertility traits. These traits varied least over PC3, which is better characterized by number of animals ranging from 28 (environment pc_1) to 130 (pc_{15}), with 57 animals in the average environment (pc_8). A detailed analysis of the phenotypic changes was given in Windig et al. (2005c). Milk production decreased not only with decreasing production intensity and relative performance but also, although less, with increasing fertility and scale (Table 8.1). Likewise, NINS decreased not only with increasing average fertility but also with decreasing production intensity and relative performance and slightly with increasing scale, while DFS increased with decreasing production intensity and slightly with increasing relative performance and decreasing scale. Changes in SCS were smallest for production intensity and largest for relative performance, decreasing with increasing intensity, fertility and relative performance and decreasing with scale.

These changes were accompanied by heterogeneous phenotypic variances (Table 8.1). Largest changes in variance were seen for DFS and NINS, which were more variable at lower production intensities, relative performance and fertility. Changes in variance were less dramatic for milk and in opposite direction from the fertility traits, while changes in variance for SCS were relatively small. Generally higher variances occurred in environments with higher means, but opposite trends in means and variances were observed for SCS and for milk when relative performance was used as the environmental variable.

Table 8.1. *Variation in mean, phenotypic variance and heritability across herd environments. Animals were ranked according to environmental values (principal components) and grouped into 15 environmental groups (pc_{env1} to pc_{env15}). μ = mean phenotypic value, σ^2 = total (phenotypic) variance. Average = value in herd environments with average environmental value (pc_{env8}), range = value in pc_{env1} – value in pc_{env15}.*

	Production intensity (PC1)		Fertility (PC2)		Scale (PC3)		Relative performance (PC4)	
	Av.	Range	Av.	Range	Av.	range	Av.	Range
305d milk								
μ	7417	6777-8159	7554	7670-7406	7670	7877-7396	7496	7135-7999
σ^2	9766	0.78-1.19	9901	1.15-1.06	10007	1.05-0.97	10018	1.10-0.91
h²	0.49	0.98-1.08	0.51	0.77-1.05	0.50	1.06-0.93	0.49	0.98-1.03
SCS								
μ	2.23	2.72-1.89	2.17	2.52-2.08	2.16	2.53-2.12	2.29	2.79-2.02
σ^2	1.261	1.01-1.11	1.286	0.99-1.05	1.319	1.00-1.02	1.278	1.02-1.01
h²	0.22	1.09-0.83	0.20	1.06-0.99	0.20	0.96-1.02	0.21	1.00-1.03
DFS								
μ	88	109-74	88	94-85	89	90-87	87	86-92
σ^2	1040	1.79-0.61	965	1.26-0.94	985	0.96-1.04	985	1.07-0.92
h²	0.09	0.87-1.01	0.09	0.92-1.06	0.08	1.09-1.00[1]	0.09	0.91-1.03[1]
NINS								
μ	2.03	1.87-2.27	2.05	2.34-1.82	2.04	2.26-2.00	2.07	2.45-1.90
σ^2	1.935	0.81-1.38	2.144	1.32-0.68	2.053	1.24-0.87	2.092	1.27-0.83
h²	0.03	0.59-1.28	0.03	1.15-0.97	0.03	0.97-1.03	0.03	0.81-1.19

[1]Variable additive genetic variance model did not converge. Changes in heritabilities are entirely due to variable residual variances over herd environments

Univariate analysis

Trends in additive genetic variances over environments (not shown) were generally similar to phenotypic variances. For DFS the REML analysis did not converge for scale and relative performance, except when a fixed additive genetic variance over environments was assumed. Heritabilities were relatively constant because ratios of phenotypic and additive genetic variances were similar over environments. Highest heritabilities were found for milk (about 50%) and for SCS (20%), while heritabilities for the fertility traits were relatively low (9% for DFS and 3% for NINS). Trends for additive genetic variances of milk were opposite to trends in its phenotypic variance when herds ranked by average fertility (PC2). Opposite trends for additive and phenotypic variances were also observed for SCS when herds were ranked by
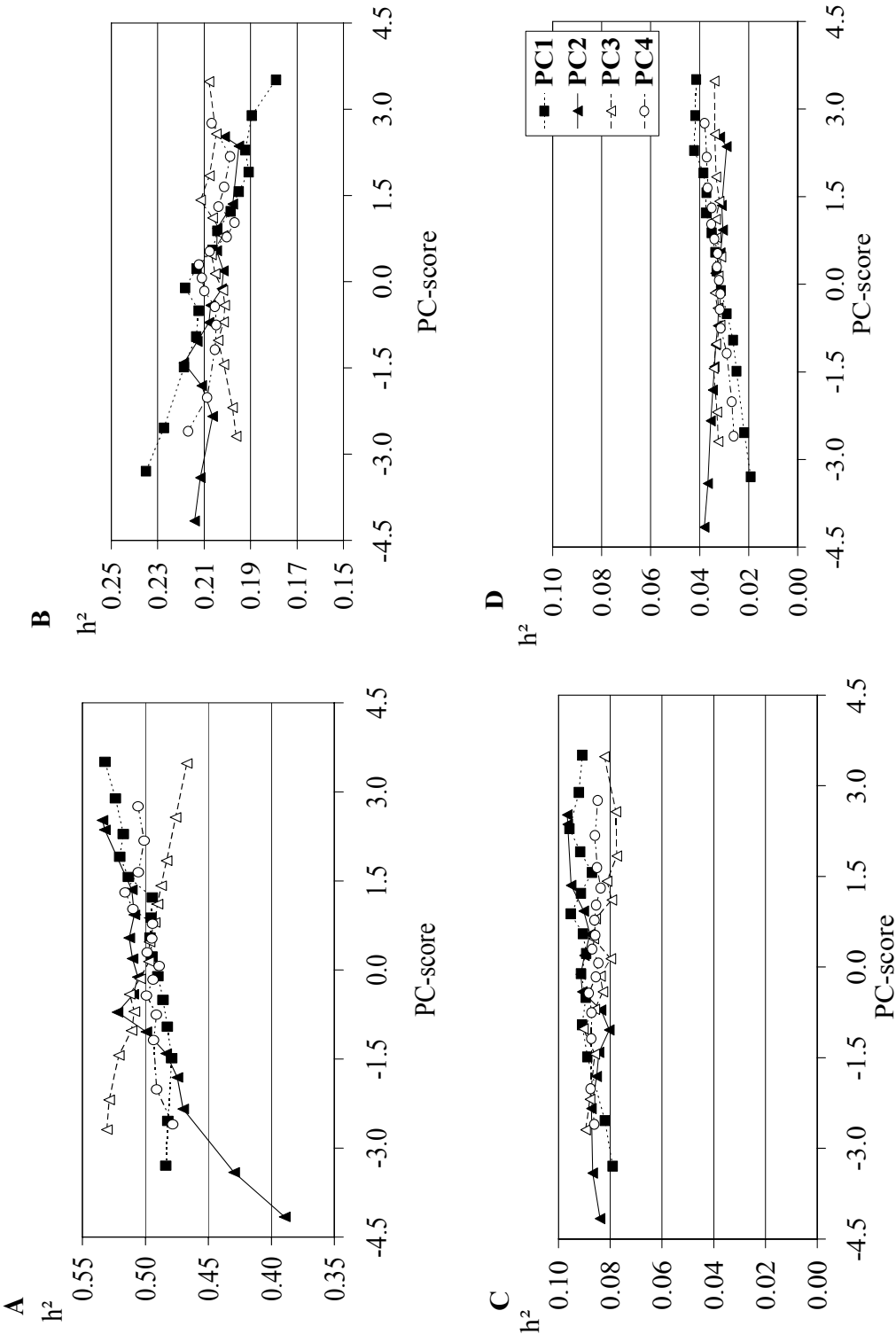
Figure 8.1. *Change in heritabilities over herd environments. Herd environments measured as four different principal components: PC1 production intensity (squares), PC2 fertility (solid triangles), PC3 scale (open triangles) and PC4 relative performance (open circles). X-axis principal components, y-axis heritabilities for : A milk, B SCS, C DFS, D NINS.*

production (PC1). Consequently the heritabilities for these two trait – PC combinations showed the largest changes (Figure 8.1), 14.5% and 5.6% respectively. Changes in heritabilities of fertility traits were smaller, but changes could be relatively large because the heritabilities themselves were smaller. Largest changes were seen for DFS – fertility, increasing from 8.4 to 9.7% and for NINS – production intensity where the heritability more than doubled from 1.9% to 4.2% with increasing intensity.

Genetic correlations within traits across environments (Table 8.2) were smaller, indicating more changes in ranking of breeding values for the two fertility traits than for milk production and SCS. For the combinations DFS – fertility, NINS – fertility and NINS – scale the genetic correlation between the extreme environments (pc_1 and pc_{15}) was less than 0.9, while it was close to 0.9 for DFS – intensity. Other genetic correlations were mostly around 0.97, except for SCS – fertility, SCS – scale and milk – relative performance which were above 0.99. Analyses for DFS - scale and DFS - relative performance did not converge, probably due to lack of variation in additive variance across environments.

Table 8.2. *Genetic correlations (approximate standard error is given as subscript), within traits, between lowest and highest analyzed environments (pc_{env1} and pc_{env15}).*

	305d Milk	SCS	DFS	NINS
Production intensity (PC1)	$0.976_{0.019}$	$0.985_{0.027}$	$0.911_{0.085}$	$0.970_{0.157}$
Fertility (PC2)	$0.960_{0.025}$	$0.997^{[1]}$	$0.883_{0.097}$	$0.899_{0.157}$
Scale (PC3)	$0.960_{0.025}$	$0.998_{0.015}$	$1.000^{[1]}$	$0.841_{0.153}$
Relative performance (PC4)	$0.998_{0.009}$	$0.951_{0.035}$	$1.000^{[1]}$	$0.974^{[2]}$

[1] Model with heterogeneous additive (co)variances did not converge
[2] Bounding of variance components prevented calculation of standard errors

Multivariate analysis

Genetic correlations between milk and SCS in the average environments varied around a mean of 0.35 while the genetic correlations between milk and DFS and milk and NINS tended to be somewhat higher (Table 8.3). These genetic correlations between the traits and milk in the same environment varied considerably over environments (Figure 8.2): milk with SCS from 0.25 to 0.47, with DFS from 0.30 to 0.48 and with NINS from 0.18 to 0.64. In some instances the genetic correlation with milk production in another environment was higher than the correlation with milk in its own environment (Table 8.3). For example, the genetic correlation between NINS and milk in low production herds was 0.21, while the genetic correlation between NINS in low production herds and milk in high production herds was 0.67.

Table 8.3. *Genetic correlations within and across environments between 305 day milk production (on each row) and fertility traits and SCS. Animals were ranked according to environmental values (principal components) and grouped into 15 environmental groups (pc_{env1} to pc_{env15}). For each Principal Component – trait combination nine correlations are given. On diagonal correlations with milk in the same environment. Top rows correlation of traits with milk in low environment (pc_1), middle rows correlation with average environment (pc_8) and bottom rows in high environment (pc_{15}).*

	SCS			DFS			NINS		
Prod. intensity - PC1	Low	Av.	High	Low	Av.	High	Low	Av.	High
Low	0.29	0.30	0.31	0.39	0.45	0.48	0.21	0.14	0.09
Average	0.34	0.35	0.35	0.34	0.38	0.41	0.47	0.40	0.34
High	0.38	0.39	0.39	0.21	0.26	0.29	0.67	0.56	0.49
Fertility - PC2	Low	Av.	High	Low	Av.	High	Low	Av.	High
Low	0.39	0.37	0.36	0.48	0.48	0.48	0.32	0.28	0.18
Average	0.35	0.34	0.33	0.33	0.37	0.39	0.53	0.47	0.43
High	0.33	0.29	0.28	0.23	0.28	0.31	0.68	0.66	0.64
Scale - PC3	Low	Av.	High	Low	Av.	High	Low	Av.	High
Low	0.35	0.37	0.40	0.30[1]	0.30[1]	0.30[1]	0.60	0.64	0.68
Average	0.33	0.35	0.38	0.35[1]	0.35[1]	0.35[1]	0.39	0.43	0.47
High	0.31	0.33	0.36	0.42[1]	0.42[1]	0.42[1]	0.12	0.14	0.18
Rel. perf. - PC4	Low	Av.	High	Low	Av.	High	Low	Av.	High
Low	0.47	0.46	0.45	0.42[1]	0.42[1]	0.42[1]	0.34	0.35	0.36
Average	0.37	0.36	0.36	0.36[1]	0.36[1]	0.36[1]	0.28	0.30	0.32
High	0.27	0.26	0.25	0.30[1]	0.30[1]	0.30[1]	0.21	0.24	0.26

Approximate standard errors for SCS on average 0.066 (ranging from 0.048 to 0.082), for DFS 0.089 (0.071-0.107) and for NINS 0.113 (0.095-0.130).
[1]Variable additive genetic (co)variance model did not converge. Variation in genetic correlations is entirely due to variable additive genetic variance over environments of milk only.

Genetic correlations decreased with increasing relative performance, but for the other PC's there were no consistent trends (Figure 8.2). The trends of the correlations of milk with DFS and milk with SCS were similar, except over production levels. The trends of the correlations of milk with DFS tended to be opposite to the trends of the correlations of milk with NINS, except over relative performance. The genetic correlations of milk with NINS were the most variable. They were weakest (<0.2) in large herds, and strongest (>0.6) in small herds and high fertility herds. Genetic correlations of SCS with milk tended to change less than correlations with DFS and NINS with milk, and were almost constant over herds ranked according to scale (PC3).
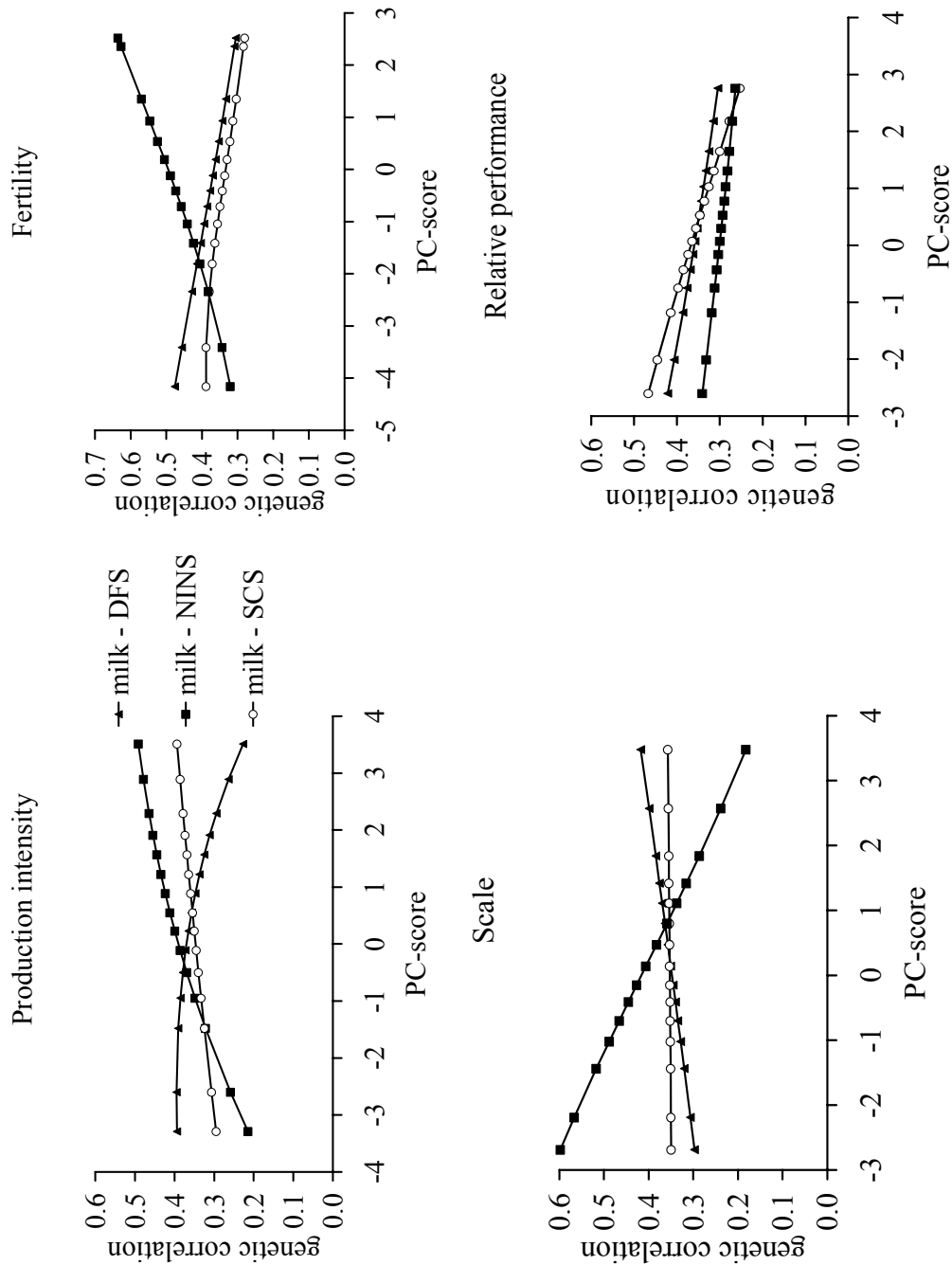
Figure 8.2. *Trends in genetic correlations between 305 day milk production and SCS (open circles), DFS (triangles) and NINS (squares). X-axis principal components A: production intensity, B: fertility, C: scale and D: relative performance, y-axis genetic correlations.*

Table 8.4. *Correlated responses to selection in the average environment for an increase in milk of 1000kg or a decrease in DFS, NINS and SCS. Animals were ranked according to environmental values (principal components) and grouped into 15 environmental groups (pc_{env1} to pc_{env15}). The average environment is pc_{env8}. Correlated responses are given for pc_{env8} itself and for pc_{env1} (=low) and pc_{env15} (=high). PC1 production intensity, PC2 fertility, PC3 scale, PC4 relative performance.*

Selection in average. environment for	Correlated response of	Response environment	PC1	PC2	PC3	PC4
		low	7.2	4.8	4.4	4.7
Milk +1000 kg	DFS	average	6.2	4.9	4.5	4.7
		high	5.1	5.1	4.5	4.6
		low	0.11	0.25	0.16	0.10
Milk +1000 kg	NINS	average	0.14	0.17	0.15	0.11
		high	0.16	0.13	0.16	0.12
		low	0.27	0.26	0.24	0.28
Milk +1000 kg	SCS	average	0.26	0.25	0.26	0.27
		high	0.25	0.24	0.28	0.26
		low	-193	-201	-183	-100
DFS: -6 days	Milk	average	-190	-168	-117	-81
		high	-147	-134	-36	-61
		low	-35	-72	-183	-100
NINS: -0.1	Milk	average	-111	-134	-117	-81
		high	-178	-198	-36	-61
		low	-176	-233	-269	-327
SCS: -0.5	Milk	average	-233	-229	-243	-247
		high	-298	-209	-214	-167

Selection response

The estimated correlated response to selection for an increase in milk production in average environments were rather constant over environments (Table 8.4). However, the response in milk production differed substantially over environments with selection for DFS, NINS and SCS in the average environment. For example, selection for a decrease in NINS of

0.1 in herds with average fertility caused a decrease in milk in the same environment of nearly 111kg. However, in herds with low fertility this decrease was only 35kg, while in herds with high fertility the decrease was 178kg. In general, selection to reduce SCS or improve fertility decreased milk production and vice versa. Selection for reduced SCS led to a decrease in milk especially in high production herds and herds with a low relative performance. Selection for shorter DFS caused a large decrease in low production and low fertility herds, while the largest decreases in milk caused by selection for less NINS occurred in high production, high fertility and small herds. On the other hand, selection for an increase in milk had an especially high correlated response for NINS in low fertility herds.

DISCUSSION

The relationship between milk yield and the fertility traits and SCS varied considerably over herd environments. Generally in dairy cattle breeding little importance is given to genotype by environment interaction. Genetic correlations in production across environments tend to be close to unity and reranking of sires is consequently rare or absent. Heterogeneity of variances is the most common effect of genotype by environment interactions observed, but methods have been developed to take these into account when estimating breeding values (e.g. Meuwissen et al., 1996). In this paper strong correlations across environments are also observed, though for the fertility traits somewhat lower values, down to 0.84, were observed. However, this may not be taken as an indication that G×E was not important. Genetic correlations between traits varied more (Figure 8.2) so that in combination with heterogeneous variances (Table 8.1) the responses to selection in an average environment differed widely across environments (Table 8.4). The implication is that it is difficult to weigh the relative importance of different traits, when breeding values are based on another environment.

Variable genetic variances and heritabilities have been reported before. The trend observed when no variable residual variances were allowed was that heritabilities of production traits increased with production levels (e.g. Hill et al., 1983; Veerkamp and Goddard, 1998; Hayes et al., 2003). This trend was also observed with heterogeneous residual variances (Kolmodin et al., 2002; Raffrenato et al., 2003). Kolmodin et al. (2002) also reported that heritabilities increased for days open with increasing herd averages for days open. Likewise, in the current study the heritability for milk increased with increasing production levels and heritability of NINS was higher in the environment with more inseminations, i.e. the low fertility environment, and in the high production environment. Heritability of DFS, however, was higher in the high fertility environment where DFS itself was lower. Over all trait – environment combinations there was no consistent trend such as, for example, that heritabilities for all traits were consistently higher in the environments where the mean was higher.

Relative performance was the fourth principal component in which average fertility and SCS were evaluated relative to average production. Thus herds where the fertility was high and SCS was low despite a high production received a high score. Whereas, herds where the fertility was low and SCS high and production at the same time was low received a low score. There was a consistent trend in genetic correlations over herds differing in relative performance. In herd environments where the relative performance was low, genetic correlations with milk were stronger, i.e. less favourable. Phenotypically, the herds with low scores for relative performance had the highest SCS levels and lowest fertility. The strong genetic correlations in these herds indicated stronger trade-offs between production and fertility and/or health. Possibly, management and genetics were not well matched in these herds. Raffrenato et al. (2003) also reported less favourable correlations in low production environments.

Apart from the genetic correlations along the relative performance axis there were no consistent trends in correlations – environment combinations. The trend in genetic correlations of milk with DFS was opposite to the trend with NINS for herd environment measured as production, fertility and scale. Kolmodin et al. (2002) reported stronger genetic correlations of days open with milk both in lower production and higher fertility herds. Furthermore, in our study was found that the trends in response to selection in the average environment generally agreed with the trends in genetic correlations with milk in the same environment.

Depending on the environment where selection takes place strong correlated responses may occur over the entire environmental range or be more restricted. For example, if selection took place in low fertility herds for an increase in milk, a relatively strong response in DFS would have occurred not only in low fertility herds but also in average and high fertility herds (Table 8.3). On the other hand while selection for a decrease in NINS in high fertility herds would have resulted in a relatively strong response in milk in high fertility herds itself, the response in low fertility herds would have been relatively weak. With selection for an increase in milk the genetic correlations indicated that NINS would have increased (i.e. fertility would have decreased) especially if selection had taken place in small herds, herds with a high fertility and herds with a high production (Table 8.3). This increase would not have only occurred in herds in the same environment, but also in large herds, herds with a low production and herds with a low fertility. As a consequence it may be interesting to change the environment in which selection takes place depending on the breeding goals. Variable genetic correlations may also influence the relative importance of traits in indices used for selection. Calus et al. (2005c) showed that the relative importance of fertility to yield traits could double across environments and that possible re-ranking based on a total merit index occurred.

From a biological viewpoint one may interpret strong genetic correlations as trade-offs between for example fertility and milk production. Trends in correlations were, however, for PC1, PC2 and PC3 opposite for NINS and DFS. These traits are clearly two different aspects of fertility from the genetic viewpoint. One possible reason is that DFS depends on the decision of the farmer when to inseminate while NINS depends on the cow. However, the farmer will base his decision more on the phenotypic value for milk production than on the genotypic value. DFS and NINS also differ from a physiological viewpoint. DFS depends on heat detection which is determined by oestrus expression levels while insemination success depends, amongst other things, on quality of embryos (pers. comm.. T. van der Lende). Oestrus levels are lower in high producing cows (Lopez et al., 2004) while embryo quality is shown to be better in non-lactating cows than in lactating cows (Sartori et al., 2002). Thus, although both DFS and NINS were negatively related to production a simple trade-off between fertility and milk production is an oversimplification.

Variation in herd environment was measured in the current study using principal components. The first 4 principal components explained about 33% of the total variance in all traits. Thus, a substantial part of the variation in the environmental variables was not explained by the PC's. For example, some soil types were not associated with high or low production, fertility or farm size. By limiting the analysis to the first four PC's, only that part of the variation in the environment that was covered by several correlated environmental variables was analyzed in this study. As a consequence an underlying environmental parameter, such as production intensity influencing several environmental variables, could be uncovered. However, one should keep in mind that other uncorrelated environmental variables can still be of interest. These variables can be analyzed as a single trait describing herd environment.

Generally, herd environments have been measured using a single trait (Calus and Veerkamp, 2003; Hayes et al., 2003), often in the form of the herd average of the trait itself that was being analyzed (Kolmodin et al., 2002). Both PCA and single environmental variables have their merits. PC's have the disadvantage that they may be difficult to interpret. In the current study the first four PC's were relatively straightforward, but higher order PC's were not. Single trait herd averages used as the environmental variable is the logical method if there is a specific question about the effect of an environmental parameter. For example, Berry et al. (2003) found that genetic variance for body condition score increased with improving silage quality. A disadvantage of single trait environmental variables is that the response might be to a correlated environmental variable instead of the variable itself, for example, the herd average of DFS was strongly correlated to the herd average of production. If there was a trend in genetic correlations over DFS herd levels this might or might not have been due to production effects. With a PCA all correlated variables are combined into new variables. In the current study DFS and milk yield along with other production or production

related variable were combined into the first PC, and DFS and other fertility or fertility related variables into the second PC. Thus with a PCA effects of overall combined effects, such as overall fertility or relative performance are evaluated.

The use of a dependent variable, such as milk production of individual animals, in the explanatory variable such as average herd milk production, is sometimes seen as a disadvantage of reaction norm models. This concern is partially alleviated by the use of PC where the explanatory variable consists of more than the dependant variable alone. Moreover, Calus et al. (2004) showed that the definition of environmental parameters including or excluding information from animals themselves in own herd averages hardly influenced estimation of genetic parameters.

Reaction norms provide the opportunity to estimate genetic parameters for an infinite number of environments. In practice, however, estimation of parameters should be restricted to the range of environments for which sufficient data are available. One reason is that if a function is extended into environments without data it is assumed that the trend (e.g. increasing variances) is the same over the whole range. A change in a trend cannot be detected if data are missing. Unless the estimated reaction norms run parallel polynomial models inevitably result in larger variances in extreme environments (Stearns et al., 1991). If this is not the case many data points in the extreme environments are required to counter this effect. This study restricted the estimation of the genetic parameters to the range of the environment of the 4500th animal to the 67,500[th] animal, with animals ranked according to the environmental values, so that enough data around these points were available for reliable estimation. The disadvantage is that one restricts the results to the less extreme environments. For example, average 305 d production along the PC1 (production intensity) varied in this paper from 6500 to 8500 kg. In more extreme environments genetic correlations might be more extreme, but a reliable estimate of these correlations cannot be provided.

For the estimation of heritabilities a second reason for restricting the environmental range for which estimations were made was that residual variances also varied over environments. Because data for a single animal are generally restricted to one environment only, one cannot estimate residual variances of reaction norm components. In this study variation in residual variances were estimated by grouping animals based on the environments in which they were measured. This can only be done with sufficiently large groups. Small group sizes result in unstable residual variances. When group sizes were halved residual variances varied from one extreme to the other over short stretches of the environmental range. Working with overlapping groups further smoothed the sudden jumping of residual variances across environments.

CONCLUSION

Genetic correlations between milk and fertility traits and SCS differed considerably over environments. Consequently the response in one trait to selection for another trait also differed over environments. Furthermore, if selection took place in one environment, but the response occurred in another environment, responses were different. It is thus important to take into consideration the environment in which breeding values and genetic correlations are determined when the effect of milk production on health and fertility traits and selection on these traits is evaluated.

ACKNOWLEDGMENTS

Chapter 9

General Discussion

M. P. L. Calus

Animal Sciences Group

P.O. Box 65, 8200 AB Lelystad, The Netherlands

GENERAL DISCUSSION

The research presented in this thesis is part of a larger research project that, in order to address societal concerns, aims at investigating health and fertility risks of high producing dairy cows. One of the questions addressed is whether, next to effects of management and genetics separately, the interaction between management and genetic level for milk yield can lead to increased health risks. The larger research project consists of an animal experiment, genetic-epidemiological research, and evaluations of opinions and direction of different stakeholders during discussion meetings.

The main objective of this thesis is to investigate the magnitude of genotype × environment interaction (G×E) for yield, health and fertility traits in dairy cattle using random regression models (RRM). After discussing the main results, the aim of the general discussion is to reflect on the methodology and the implications for animal breeding. Furthermore, the results will be linked to broader issues of robustness of dairy cows, risks of high milk yield, and the farmers' attitude towards G×E for health and fertility.

GENOTYPE × ENVIRONMENT INTERACTION FOR YIELD, HEALTH, AND FERTILITY TRAITS

In this thesis the investigation of G×E was approached by estimation of variance components to identify one or more of the following effects of G×E: heterogeneous variances across environments, genetic correlation of a trait expressed in different environments being less than 1.0 (reranking), and heterogeneous genetic correlations between traits across environments. Those effects were estimated for yield (chapter 3), and health and fertility traits (chapters 5, 6, 7, and 8), based on a large number of herd characteristics, chosen to reflect herd environment and management style of the farmer. Of all these estimates, significant G×E was detected in 86% of the situations for yield traits, but only in 14% of the situations for health and fertility traits. This suggests that G×E is relatively unimportant for health and fertility traits. However, in chapter 4 it was found that the power to detect genetic variance for the slope of a linear reaction norm is lower for low heritability traits than for high heritability traits. In the cases where G×E was found for health and fertility, large differences in genetic variances were observed across environments: genetic variances for fertility traits increased in some situations more than twofold, and a threefold increase for genetic variances of SCS was found (chapter 7). For yield, the variances at most doubled across environments (chapter 3). Genetic correlations of a trait across environments were as low as 0.65 for survival (chapter 5), while for SCS, on a test-day level, the lowest genetic correlation was as low as 0.72 (chapter 7). Also, in the literature (chapter 2) within country genetic correlations for health and fertility were reported as low as 0.74. Genetic correlations between yield traits across environments were all close to unity (chapter 3). Further, in chapter 8 it was demonstrated that the magnitude of genetic correlations between yield and health and fertility differed

across environments (albeit in all herd environments the correlations were antagonistic). These specific effects might be important, but the general conclusion might be that estimated G×E effects mainly consisted of heterogeneous genetic variances with limited reranking.

Of the many environmental parameters (EP) describing herd environment, EP based on the phenotypic herd average of the trait analyzed, appeared to be associated with G×E effects, especially in terms of heterogeneous variances. Furthermore, herd-year average fat-to-protein ratio, change in fat percentage between 14 and 77 DIM, and body condition score were associated with heterogeneous variances and reranking for several health and fertility traits (chapter 5). For yield traits, the same EP were associated with no G×E (change in fat percentage between 14 and 77 DIM), limited heterogeneous variances without reranking (fat-to-protein ratio), or relatively large heterogeneous variances without reranking (body condition score) (chapter 3).

Lower values for change in fat percentage during the first part of the lactation are associated with longer lasting and more severe negative energy balance (De Vries and Veerkamp, 2000). Lower fat-to-protein ratio is associated with higher proportions of concentrate and less fiber in the diet (Bargo et al., 2003) and higher energy balance of the cow (Grieve et al., 1986). Condition score is indicative of the energy balance of the cow as well (Veerkamp, 1998). Many environmental parameters were investigated in this thesis, such as production level, farm size, average somatic cell score and calving interval, however, it appears that the herd parameters linked to nutrition and energy balance are most important for G×E, although no direct measures of these factors were available in this study.

METHODOLOGY – REACTION NORM MODELS
Dimensions of the model

Reaction norm models have been used to estimate G×E throughout this thesis. Covariance functions (Kirkpatrick and Heckman, 1989), estimated by random regression, were used to model genetic effects as function of a continuous environmental parameter. Another approach that has been used widely to estimate G×E is a multi trait model (also known as character state model). In this model, herd environments are grouped and the same trait in another group of environments is considered to be a different correlated trait. Reaction norm models were chosen above multi trait models because arbitrary grouping of environments is avoided, and the number of estimated variance components can be more parsimonious.

The models fitted were expanded from the most simple reaction norm model with a linear random regression on one continuous environmental parameter, to higher order random regressions (chapters 3, 5, 6, and 7), to models including two parameters to define the environment (chapter 7), and to multivariate reaction norm models (chapter 8). Higher order random regressions up to the third order had a significantly better fit for yield traits (chapter 3) than lower order regressions, but only in a few cases second order random regressions were

required for health and fertility traits (chapter 5). The differences in order of fit of the random regressions for yield and health and fertility may to some extent be caused by the differences in heritabilities of the traits.

In chapter 7, one environmental parameter reflected herd environment (bulk milk somatic cell count), while the other represented the 'cow environment' (days in milk). The interaction of bulk milk somatic cell count and days in milk was included as a third environmental parameter, to account for different levels of G×E at different stages of lactation. Similarly, Veerkamp and Goddard (1998) showed that genetic parameters for yield traits are influenced not only by month of lactation and herd production separately, but also by an interaction between those two factors. In both situations, including a random regression on stage of lactation and on the interaction between stage of lactation and the environment, allowed more detailed analysis of phenotypic information, e.g., for different stages of lactation rather than for lactation averages. Comparing the genetic correlation for SCS across environments based on lactation records (i.e. 0.97 based on heifer and 0.93 based on cow data; chapter 5) to the genetic correlation across environments based on test-day records (i.e. 0.72 early vs. 0.92 late in lactation; chapter 7), indicates that more G×E can be revealed when estimated from more detailed phenotypic information.

Alternative models to estimate G×E

As well as reaction norm models and multi trait models, other models could have been used to estimate G×E, such as character process models, structured antedependence (SAD) models (Nunez-Anton and Zimmerman, 2000), and splines. At the start of this thesis, little experience existed with those alternative models, but in the mean time more experience has been gained with those models. Although grouping of environments is not necessarily avoided in those models, the number of estimated parameters may be smaller than for the RRM and the multi trait model. The difference in variance structure may give higher flexibility to for instance SAD models compared to RRM (Nunez-Anton and Zimmerman, 2000; Jaffrezic et al., 2004). Therefore, the use of RRM and SAD models to estimate G×E was compared.

In a simulation study, a first order structured antedependence model (SAD) and a first order RRM were compared in their ability to estimate G×E (Calus et al., 2005a). In the simulated data sets, G×E was modeled either following a RRM or structured variances. In the case of structured variances, five discrete groups of environments were formed based on increasing simulated herd effects. For each group of herds, a different genetic variance was simulated.

When G×E was simulated using structured variances, estimated sire variances of the SAD model were closer to the simulated variances than those of the RRM (Figure 9.1). Genetic correlations were overestimated by both models, but the overestimation was larger for the

RRM than for the SAD (Table 9.1). Correlations between simulated and estimated breeding values were slightly higher for the SAD model than for the RRM (Table 9.2).

When G×E was simulated following a RRM, estimated sire variances across environments were close to the simulated variances for both models. Genetic correlations were closer to the simulated values with the RRM and overestimated with the SAD model (Table 9.1). Correlations between simulated and estimated breeding values were hardly different between the SAD model and RRM (Table 9.2).

Based on the estimated genetic correlations, both the RRM (consistent with results from chapter 4), and the SAD underestimated G×E. Based on the correlations between simulated and estimated breeding values, SAD models seem to predict breeding values slightly more accurate than RRM. However, the differences were so small that both models seem equally able to estimate G×E.
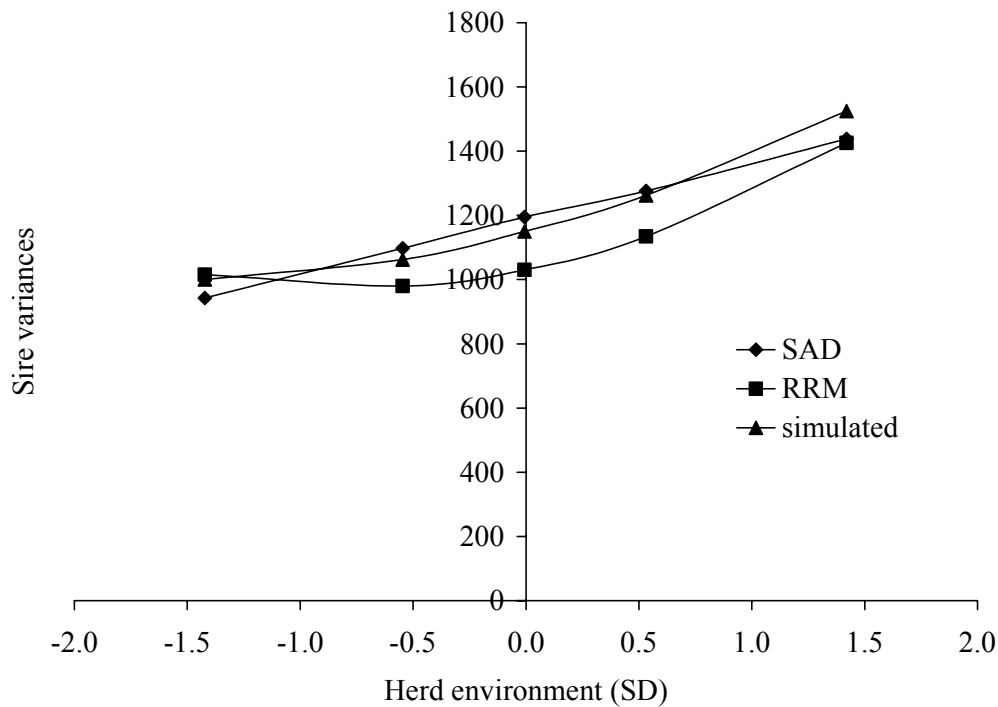

Figure 9.1. *Sire variances across environments estimated with a RRM and SAD model, based on simulations with structured sire variances across environments. The triangles mark the estimates in each of the five discrete environments for the SAD model.*

Table 9.1. *Simulated and estimated (with RRM or SAD) genetic correlations between groups of environments, based on simulations with structured variances or variances based on a random regression (SD ranged from 0.001 to 0.042).*

Environment		Structured variance			Random regression		
1	2	Simulated	RRM	SAD	Simulated	RRM	SAD
2	1	0.90	0.96	0.89	0.93	0.93	0.92
3	1	0.81	0.91	0.84	0.85	0.84	0.88
3	2	0.90	0.99	0.94	0.98	0.98	0.96
4	1	0.73	0.83	0.79	0.77	0.77	0.86
4	2	0.81	0.95	0.89	0.95	0.95	0.94
4	3	0.90	0.99	0.95	0.99	0.99	0.98
5	1	0.66	0.70	0.72	0.66	0.66	0.85
5	2	0.73	0.86	0.81	0.89	0.89	0.92
5	3	0.81	0.93	0.86	0.96	0.96	0.96
5	4	0.90	0.98	0.91	0.99	0.99	0.98

Table 9.2. *Correlations between simulated and estimated breeding values (with RRM and SAD models) in each of the five groups of environments (SD ranged from 0.001 to 0.032).*

Environment	Structured variance		Random regression	
	RRM	SAD	RRM	SAD
1	0.96	0.97	0.98	0.98
2	0.95	0.97	0.99	0.99
3	0.95	0.97	1.00	0.99
4	0.96	0.98	1.00	0.99
5	0.96	0.98	0.99	0.99

IMPLICATIONS OF ESTIMATED G×E

Breeding value estimation

In this thesis, G×E effects mainly consisted of heterogeneous genetic variances across environments. Little evidence was found for heterogeneous heritabilities across environments, apart from the heritability of SCS across DIM and BMSCC. Heterogeneous heritabilities across environments can cause differences in accuracy of testing animals in different environments (Hill et al., 1983), as shown for estimated breeding values for SCS depending on DIM and BMSCC (chapter 7). A consequence of heterogeneous genetic variances is that animals from environments with larger variance are favored when this form of G×E is ignored during selection (Hill et al., 1983). To overcome this selection bias, breeding values are usually estimated using a correction for heterogeneity of phenotypic variances (e.g.,

134

Meuwissen et al., 1996). These corrections 'scale' estimated breeding values to an average environment.

Another consequence of heterogeneous variance is that selection (e.g. using estimated breeding values for an average environment) leads to different absolute responses in different environments. For example, it was shown that selection for fertility will yield higher selection responses in herds with on average poorer fertility (chapter 5), and that selection for reduced SCS will yield higher selection responses in herds with on average higher SCS (chapter 7). The models described in this thesis, incorporate heterogeneous variances across environments in estimated genetic variances and breeding values, as well as genetic correlations smaller than 1.0. As a result, a specific ranking and scale of estimated breeding values can be produced for every herd environment. However, for traits that have genetic correlations across environments close to unity, use of a correction for heterogeneity of variances in breeding value estimation is likely to be sufficient, as ranking on breeding values for a single trait is not affected by heterogeneous variances. For traits that have genetic correlations across environments that are considerably lower than unity, use of reaction norm models would be necessary to accurately predict herd-specific breeding values. Based on the results in this thesis, it can be concluded that in breeding value estimation on a single trait basis, correction for heterogeneity of variances is sufficient for nearly all traits, and that one list of breeding values identifies the best sires for all environments.

Breeding goal

The discussion whether to present one list of breeding values (after adjustment for heterogeneity of variances) or to present environment specific breeding values is not only affected by single trait breeding values estimation. Heterogeneous variances of different traits, even with no reranking across environments for the individual traits, can cause reranking across environments based on a total merit index (Namkoong, 1985). This could be considered as a fourth form of GE: the correlations between total merit indexes across environments are smaller than 1.0. It was shown that the relative importance of fertility, based on genetic standard deviations, compared to protein yield was twice as large in a 'poor' fertility environment compared to a 'good' fertility environment (chapter 5). Furthermore, it has been reported that the economic value of fertility (say a decrease of DFS of one day) is larger for animals with poor fertility than for animals with good fertility (Esslemont et al., 2001). The higher genetic variance combined with the larger economic value for DFS in a poor fertility environment, would lead to an even greater G×E effect. This seems to contradict the common believe that in herds with poor fertility focus should be on management improvement and not on selection.

The breeding goal is defined as the sum of the expected response (i.e. breeding values) in the relevant traits, multiplied with their respective economic values. Therefore, when

estimating G×E effects on a breeding goal level, not only the change in genetic parameters across environments should be taken into consideration, but also differences in economic values across environments. A framework with both genetic parameters and economic values depending on herd environment enables derivation of total merit indexes for all possible environments, using information from all possible environments while accounting for G×E (Kirkpatrick and Bataillon, 1999). Different strategies can be followed for such a framework. The most sophisticated option would include herd-specific breeding values and herd-specific economic values. More practical options involve one or a limited number of bull rankings, in which differences in breeding values and economic values across environments are accounted for. When the effect of G×E on breeding values only exists of heterogeneous genetic variances, and heritabilities are homogeneous across environments, a correction for heterogeneity of variances can be used to account for G×E. When economic values are different across environments, an average economic value across environments can be calculated, while accounting for the frequency of herds with different economic values. In this way, the breeding goal is implicitly defined for an 'average' environment. Based on results in this thesis, for most traits accounting for heterogeneity of variances seems sufficient. In order to predict the outcome of selection for specific herd environments, the 'average' breeding values can be translated to herd-specific breeding values using herd-specific genetic variances. Heterogeneity of variances can be accounted for in economic values (as shown in chapter 5), that conveniently can be used in mating programs to select bulls on a herd level (Bowman et al., 1996).

Breeding programs

When the correlation between breeding goals is close to unity, collaboration of breeding programs across environments is beneficial, because the sampled population increases, resulting in higher selection intensity and higher genetic gain. Presence of reranking might reduce benefits of collaboration of breeding programs across environments, and in extreme cases, collaboration might actually not lead to higher genetic gain (Mulder and Bijma, 2005). For which of the traits collaboration between breeding programs across environments would be recommended, can be anticipated from the estimated genetic parameters in this thesis and in the literature. For yield and fertility traits, estimated genetic correlations across environments were close to unity, and for yield between countries the same was reported to be the case (Mark, 2004). These high genetic correlations imply, that selection for yield and fertility across environments within countries, and for yield between countries, is beneficial and that collaboration of breeding programs will lead to higher genetic gain (Mulder and Bijma, 2005). For SCS, mastitis, and survival, in some cases estimated genetic correlations across environments were between 0.6 and 0.9, and reported genetic correlations between countries had comparable values (Mark, 2004). These lower genetic correlations imply that

collaboration between breeding programs in different environments will not lead to higher genetic gain for SCS, mastitis, and survival (Mulder and Bijma, 2005).

ROBUSTNESS OF DAIRY COWS

Environmental sensitivity of dairy cows is often associated with recent issues such as the more negative energy balance of high genetic merit cows and the diversification of husbandry systems. These issues stimulate the call for 'more robust' dairy cows. One way of describing robustness is lack of environmental sensitivity: robust animals are not easily affected by changes in the environment. Environmental sensitivity is mathematically represented by the slope of an animals reaction norm (Falconer, 1990). This provides a direct link to G×E on the population level: differences in environmental sensitivity (i.e. robustness) between genotypes will show as G×E. Intuitively, environmentally sensitive animals (i.e. with poor robustness) are believed to require a stable environment that supports high performance, in order to perform optimally. Some studies indicate that in a continuously improving environment, selection for increased performance leads to increased environmental sensitivity for the trait selected on (i.e., reduced robustness) (Falconer, 1990; Kolmodin et al., 2003; Van der Waaij et al., 2004). Based on results in this thesis, here it will be discussed whether environmentally sensitive or insensitive animals are expected to be selected, and thus what the effect is of selection on environmental sensitivity. Situations will be considered where the environment is defined as the average of the analyzed (or a closely related) trait following Falconer (1990), as in these cases the average reaction norm of the (base) population is an increasing line. At the same time the estimated breeding values across environments give insight in how management changes and selection in the same direction together influence environmental sensitivity for a trait. The expected performance of an animal across environments is represented by the population average reaction norm plus its own breeding value, which may increase or decrease across environments. Thus, the environmental sensitivity of an animal is equal to the slope of the average reaction norm plus the slope of its own breeding value across environments.

Generally it is observed that when the environment was defined as average of the analyzed trait, the estimated genetic variance increased with increasing mean. In this thesis and in the literature this was true for yield (e.g., chapter 3, Hill et al., 1983; Boldman and Freeman, 1990; Kolmodin et al., 2002), and fertility and health traits (e.g., Kolmodin et al., 2002; chapter 5). For yield, the animals with the highest breeding values generally also had the steepest reaction norm (based on results from chapter 3). This is to be expected when the analyzed trait is used to define the environment, as it is unlikely that performance decreases in a situation with increasing environmental circumstances and no reranking. Therefore, for yield the same animals will be selected in different environments, and, as indicated by Falconer (1990), it is expected that selection for yield will favor the most environmentally

sensitive animals, irrespective of the environment of selection. For days to first service and days to last service, across herd-year average calving interval, the animals with the highest breeding values generally also had the steepest reaction norm (chapter 5). However, for days to first and days to last service, the 'best' animals are those with the lowest breeding values, and are consequently the least environmentally sensitive animals. As a consequence, selection for improved fertility is expected to result in selection of relatively environmentally insensitive and therefore robust animals.

However care should be taken when interpreting the results in terms of the effect of selection on environmental sensitivity and linking it directly to robustness. It was observed that with an increase in mean, the variance increased as well. The question is to what extent the increased genetic variance is explained by scaling effects. Here, the term scaling effect is used in the sense that with a change of the mean for a trait, the variance of that trait is expected to change as well. Scaling effects can be removed by transforming the data, although it is not as simple as turning around the measurement scale (e.g. subtracting all environmental values from the maximum value for the environmental parameter). However, the expressed genetic variance for yield will still be larger in herds with higher average yield, and the expressed genetic variance for fertility will still be larger in herds with poorer average fertility. A more appropriate measure for the variance relative to the mean is the coefficient of variation (CV). For instance, for milk yield the CV was found to be rather constant across different levels of the mean (Hill et al., 1983), while the variance for yield did increase with an increasing mean. Based on data used in chapter 3, the CV was calculated for protein yield as the phenotypic SD divided by the phenotypic mean, for the 20% herds with the lowest and highest average protein production. For herds with low average protein production, the CV was higher (i.e. 14.3%) than for herds with high average protein production (i.e. 12.3%). When the CV is roughly constant across different means, the phenotypic variance across different means can be standardized by log-transforming the records analyzed (Lynch and Walsh, 1998). Re-analyzing the data for protein yield, after log-transformation, showed that with increasing mean for protein, the genetic variance for protein hardly changed across environments. For herds with low average days to first service (DFS), the CV for DFS was lower (i.e. 31.4%) than for herds with high average DFS (i.e. 39.7%). Re-analyzing the data for DFS, after log-transformation, showed that with increasing mean for calving interval the genetic variance for DFS still increased. In conclusion, the results after the log-transformation suggest that there are hardly any genetic differences in environmental sensitivity for protein that are not associated with scaling effects. However, genetic differences in environmental sensitivity for DFS that are not associated with scaling effects seem higher in herds with high average calving interval.

Thus, when making no distinction in environmental sensitivity caused by scaling or other effects, selection for yield is expected to increase environmental sensitivity (ES) for yield, and

selection for fertility is expected to decrease ES for fertility. Whether this is desired or not depends on several factors. With respect to yield, the selected animals are the least robust animals, i.e. their actual yield is most easily affected by changes in the environment. This might be perceived as an undesirable characteristic, when the environment becomes 'less supportive' (i.e., the management change is translated into a decrease in performance), but actually as a desirable characteristic, when the environment becomes 'more supportive' (i.e., the management change is translated into an increase in performance). The increase in ES for yield traits, as a consequence of selection for increased performance, does not seem to be a problem for the actual yields, because the selected animals are expected to have the highest yield in nearly each environment. Furthermore, adjusting performance to a less supportive environment for instance as a way to match energy intake and output, might actually be seen as a 'robust' response on a different level (i.e., the overall welfare of the animal). In other words: environmental sensitivity in one trait (yield) may lead to robustness in another trait (welfare). However, as a result of the increase in ES for yield traits 1) a (temporal) change in herd environment might lead to a larger reduction in yield, which might be perceived as a problem and lead to (unnecessary) management changes or individual treatment, and 2) increased ES for yield might be associated with increased ES for other traits, such as number of inseminations per service period (chapter 8). Therefore, ignoring fertility, whilst selecting for yield, gives cows that are more environmentally sensitive for both yield and fertility, with on average poorer fertility.

ANIMAL HEALTH RISKS ASSOCIATED WITH HIGH MILK YIELD

The research presented in this thesis was part of a larger research project that, in order to address societal concerns, aimed at investigating health risks of high producing dairy cows. Earlier research has shown that genetic selection for yield alone had little risk for acute health problems in the average environment, as long as selection was performed on an index combining yield with health and fertility traits. However, the remaining question was whether the interaction between management and genetic level for milk yield could lead to increased health risks. The larger research project contained an animal experiment (Beerda et al., 2005; Ouweltjes et al., 2005), genetic-epidemiological research (this thesis, Windig et al., 2005a; Windig et al., 2005b; Windig et al., 2005c), and evaluations of opinions and direction of different stakeholders during discussion meetings. In this final part of this thesis, results of these other parts of the research project are discussed together with the findings in this thesis.

Phenotypically, increased yield caused by high genetic merit was associated with impaired fertility and udder health (Ouweltjes et al., 2005). In this thesis, lower average fat-to-protein ratio (possibly reflecting higher energy status of the cows) was related to higher incidences of predicted mastitis and poorer insemination success (chapter 5). In the animal experiment, higher energy density of the ration was associated with increased yield, better energy balance

and higher metabolism, but no effects on fertility or udder health were found (Ouweltjes et al., 2005). Higher frequency of milking was associated with increased yield, poorer energy balance, lower somatic cell count and slightly later onset of oestrus. In the genetic-epidemiological research it was shown that management was able to alter the relation between yield and udder health at herd level, but not at an individual animal level: herds with higher yield had on average better udder health whereas within herds animals with higher yield had on average poorer udder health (chapter 5, Windig et al., 2005c). This was however not the case for fertility: herds with high yields had on average poorer fertility (chapter 5, Windig et al., 2005c).

In this thesis, it was discussed that some estimated G×E effects were associated with environmental descriptors reflecting energy status of the cow. This indicates that energy status of the cow, next to direct effects on health and fertility, can also affect health and fertility through G×E. Further, genetic correlations between yield and fertility and udder health were affected by herd environment, but were antagonistic across all environments (chapter 8). The genetic correlation between milk and SCS was slightly higher for herds with higher production (as indicated by principal component 1; chapter 8). Phenotypically, risks with respect to SCS were higher for animals with higher milk yield, and especially for those that were producing in herds with high average production. With increasing herd production, the genetic correlation between milk and number of inseminations increased with increasing herd production, while the genetic correlation between milk and days to first service decreased. Phenotypically, animals with high milk yield had increasingly poorer fertility in herds with higher production (i.e. more inseminations were needed).

In conclusion, risks of high milk yield for health and fertility depend on the factors that cause the high milk yield. Higher energy density of the ration was associated with better energy balance, but at the same time higher metabolism, possibly leading to higher metabolic stress in the cows. Higher frequency of milking was associated with better SCS, but poorer energy balance and slightly poorer fertility. Therefore, avoiding risks of high milk yield evolving from management does not seem straightforward, and likely depends on the attitude of the farmer and the constraints in the herd environment. Risks evolving from high genetic merit for yield can be counteracted by applying multitrait selection for yield, health and fertility simultaneously.

FARMERS' ATTITUDE TOWARDS G×E FOR HEALTH AND FERTILITY

An important part of the research project was to evaluate which risks of high milk yield are identified by dairy farmers, how they deal with them, and how they possibly would make use of the results of the research project. One of the main aims was to evaluate whether dairy farmers think that risks of high production arise from management, breeding, or possible

		Herd management	
		Low input: Low costs Little time Low tolerance of problems	**High-tech:** Top quality feed Intensive care Fine regulation
Genotype (Breeding goal)	**High:** Yield Dairy type	12%	37%
	Less high: Robust Longevity	38%	13%

Figure 9.2. *Indication of preferences of farmers with regard to their breeding goal ("high" or "less high") and herd management ("low input" or "high-tech"), as observed during four discussion meetings with dairy farmers in different regions of the Netherlands (170 people, mainly farmers, were present at these discussions, being ~ 0.5% of all Dutch dairy farmers).*

interactions between management and breeding. This evaluation was done during four discussion meetings with dairy farmers in different parts of The Netherlands, next to discussions with representatives of different groups of stakeholders (i.e. animal rights organizations (Dierenbescherming), Dutch Ministry of Agriculture, Nature Management and Fisheries, Faculty of Veterinary Medicine, Royal Dutch Society for Animal Medicine (KNMvD), Farmers unions (LTO, NMV), Dutch Dairy Board (PZ), the cattle improvement organization NRS, and the breeding organization HG). With regard to breeding decisions, the farmers were asked whether they aimed at high yield or at robustness and longevity. With regard to management they were asked whether they aimed at low costs and low labor requirements, or at "fine tuning" and top quality feed. The results indicated that most farmers chose a combination of low input management and robust animals, or a combination of high-tech management and high breeding values (Figure 9.2.). Some farmers did choose to combine high breeding values with low input management, or robust animals with high tech management (Figure 9.2.). These results are interesting, because farmers are often described as having one goal (i.e. to increase yield), while these results clearly suggest otherwise. The choices of the farmers reflected which strategy they expected to be optimal for their own farm. Most of the farmers indicated that management and breeding should go hand in hand, whereas a reasonable proportion of the farmers indicated that negative side effects for selection on milk yield can be counteracted by improving management. The first statement is
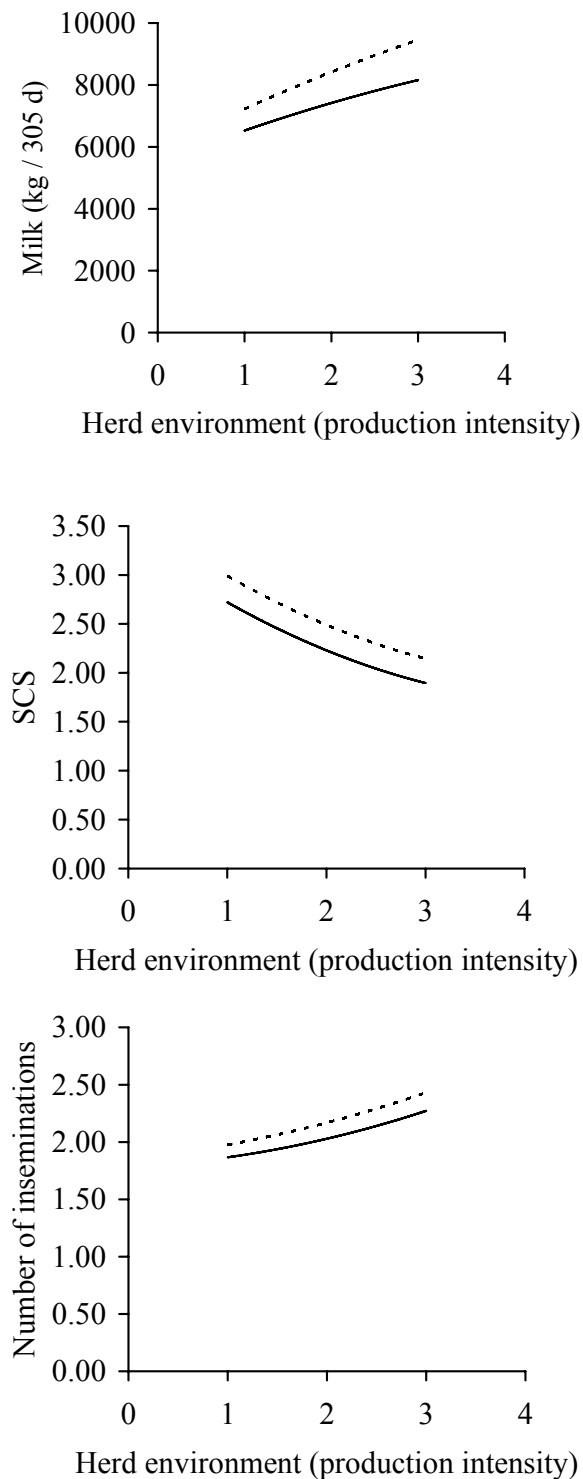
Figure 9.3. *Averages for 305 d milk production, SCS, and number of inseminations per service period across production intensity environments, and the (correlated) responses to selection for +1000 kg milk / 305 d in the average environment (given as dotted lines; based on results presented in chapter 8).*

supported by Figure 7.3 that showed that SCS can be most effectively decreased by management and breeding decisions simultaneously. To challenge the view that negative side effects for selection on milk yield can be counteracted by improving management, average 305 d milk yield, SCS, and number of inseminations, were plotted against herd environment, defined as production intensity (Figure 9.3). Higher production intensity can be interpreted as an 'increase in management to stimulate higher individual milk yield'. For each of the traits, the correlated response to selection for increased milk yield in the average environment by 1000 kg / 305 d was plotted as a dotted line. These figures indicate five things: 1) higher average production is on average related to better SCS and poorer fertility, 2) selection for higher milk yield alone leads to higher SCS and poorer fertility, 3) selection for higher milk yield alone leads to increased ES for yield and number of inseminations, but slightly decreased ES for SCS, 4) simultaneous selection for yield and improvement of management (with respect to yield) leads to poorer fertility and possibly to higher SCS, and 5) selection for yield combined with "less supportive" management (i.e., the cows are not "pushed to fully express their genetic potential"), in order to improve fertility, will lead to higher SCS. These findings indicate that selection for yield combined with management changes to improve yield, lead to poorer fertility and possibly higher SCS, while selection for yield combined with management changes to reduce yield, lead to better fertility and higher SCS. Therefore, it is concluded that improvement of management (with respect to yield) alone will not be able to counteract influence of one-sided selection for yield on health and fertility.

Summary

Summary

Genotype × environment interaction (G×E), also known as environmental sensitivity of genetic merit, is the phenomenon that different genotypes respond differently to changes in the environment. G×E can consist of the following effects: heterogeneous variances across environments, genetic correlation of a trait expressed in different environments being smaller than 1.0 (reranking), and heterogeneous genetic correlations between traits across environments. Traditionally, re-ranking of animals is often considered to be of more concern for animal breeders than the change in variance across environments, and for yield traits relatively little reranking across environments was observed. However, the recent development of including health and fertility traits in a large number of national breeding goals renews the interest in G×E. Firstly because little is known about reranking of animals for health and fertility traits across environments, and secondly because the change in variances across environments becomes important. For example, the relative importance of traits in multitrait selection is affected by changes in variances across environments, and the change in variances can lead to more phenotypes that perform under a minimum threshold, which is especially important for health and fertility traits (for more details, see chapter 2). More general reasons why there is an interest in G×E are the debates on risks of high milk yields and robustness of animals. Societal concerns exist whether milk yield can be too high and impair health of cows. As breeding is responsible for about half of the increase in yield in the last decades, another question is whether this strong selection for yield has led to selection of less robust animals, and whether the selected animals are the optimal choice for the range of different herd environments that exists today.

A recent development in the estimation of G×E is the use of covariance functions, that estimate genetic variances as a function of environmental descriptors (Veerkamp and Goddard, 1998; Calus et al., 2002; Kolmodin et al., 2002; Fikse et al., 2003b). These covariance functions can be directly estimated from the data using random regression models (Van der Werf et al., 1998), and for the purpose of estimating G×E are called reaction norm models (for more details, see chapter 2).

The main objective of this thesis is to investigate the magnitude of genotype × environment interaction (G×E) for yield, health and fertility traits in dairy cattle using random regression models.

At first a reaction norm model was applied to 14 environmental parameters, calculated as herd-year phenotypic averages, to estimate G×E for milk, fat, and protein yield in Dutch dairy cattle (chapter 3). The yield traits showed G×E in combination with nearly all environmental parameters. Hardly any reranking of sires was found, but genetic variances changed considerably and even doubled across environments in some situations. The largest variances for the yield traits were found in herds with on average high protein, high persistency, young age at calving, high body condition score, short calving intervals, and calving peak in the

autumn or winter. Reranking of sires based on an economic index combining yield traits (i.e. Inet) was limited due to comparable changes in genetic variance of the yield traits across environments.

Estimated G×E for protein yield was largest when associated to herd-year average of protein yield. In this situation the environmental parameter was calculated from the phenotypic records that were used in the analysis. This result added to the concern about possible confounding between the environmental parameter and the analyzed records. Thus, the aim of chapter 4 was to investigate effects of calculating environmental parameters from the records that are actually analysed. A simulation study was performed where traits were defined as linear functions of the simulated herd effect and both random and non-random mating of animals was considered. A proposed alternative model iteratively replaced within-herd average phenotypic performance as environmental parameter, by estimated fixed herd effects. The alternative model, however, was not able to better disentangle genetic and environmental effects. Non random use of sires, poor genetic connectedness and small herd size had a large impact on estimated covariance functions, calculated environmental parameters, and expected breeding values. The bias was such that based on unbalanced data reaction norm models underestimated G×E. This bias can be reduced using large numbers of animals per herd to calculate environmental parameters. If herd-years are small, they could be joined per herd across years. However, if these groups are stretched out over too long a time, information on G×E (i.e., the parameters explaining the environment), might become less specific and some information to estimate G×E might be lost.

In chapter 5, the association of herd environment with phenotypic levels and breeding values of fertility and health traits was investigated. Herds with lower average somatic cell score had in general more desirable values for almost all analyzed traits (i.e. days to first service was 7 days shorter), as did herds with lower average calving interval (i.e. 2.8% lower incidence of predicted mastitis). Herds with higher average protein production had slightly poorer fertility, but more desirable values for all other analyzed traits (i.e. 5.1% less predicted mastitis, 0.4 lower somatic cell score and 0.6 higher body condition score). Genetic variances varied generally only slightly across environments, but in some cases increased more than twofold across environments. The lowest estimated genetic correlations based on the heifer data were 0.76 (SE 0.21) for first service conception between herds with differing average body condition score, and based on the cow data 0.65 (SE 0.10) for survival between herds with differing average age at calving. The relative importance of days to first service compared to protein yield, measured in genetic standard deviations, was one and a half times as large in herds with an average calving interval of 410 compared to herds with an average calving interval of 370 days. Therefore, response to selection for fertility is expected to be highest in herds with poor fertility.

Summary

To estimate G×E based on more detailed phenotypic information, analysis were performed for fat yield and fat percentage (chapter 6) and somatic cell score (chapter 7) based on test-day records rather than on lactation averages. In chapter 6, G×E was estimated for fat yield and fat percentage, using a random regression on average herd test-day fat percentage, for Australian dairy cattle. The hypothesis was that G×E for fat yield and fat percentage is an indicator for susceptibility to the metabolic disorder milk fat depression. Milk fat depression was defined as 1) variation of milk fat percentage of animals within lactation, and 2) the deviation of an animal's fat percentage on a test-day from its expected fat percentage based on fat percentage on the first test-day. These traits had estimated heritabilities of 4 and 5% and genetic correlations between environments with low and high average fat percentage of 0.43 and 0.50. Genetic correlations between fat yield expressed in different environments ranged from 0.83 to 1.00. Genetic correlations between fat percentage expressed in different environments ranged from 0.87 to 1.00. Results suggested that genetic variation in susceptibility to milk fat depression is present and that selection for reduced susceptibility to milk fat depression is possible. Low susceptibility for milk fat depression was associated with a small decline in fat yield between test-days with high and low average fat percentage. Hence, low susceptibility for milk fat depression was associated with environmental insensitivity for fat yield.

In chapter 7, G×E was estimated for somatic cell score (SCS) based on test-day records of Dutch dairy cattle. The reaction norm model was further extended by not only considering a random regression on an environmental parameter (bulk milk somatic cell count) but also on days in milk and on the interaction between bulk milk somatic cell count and days in milk. This allowed individual differences in estimated lactation curves, but also differences in estimated G×E at different stages of lactation. Estimated sire variances for SCS were highest early in lactation at high levels of bulk milk somatic cell count and lowest early in lactation at low levels of bulk milk somatic cell count. Genetic correlations between SCS at the same stages of lactation, across levels of bulk milk somatic cell count, were between 0.72 and unity. The lowest correlated responses across bulk milk somatic cell count and days in milk were less than half the direct response to selection in the response environment. Responses to selection were reasonably high among environments in the second half of the lactation, while responses to selection between environments early and late in lactation tended to be low. Selection for reduced SCS yielded the highest direct response early in lactation at high BMSCC. The results indicated that using more detailed phenotypic information, i.e., on a test-day level, reveals greater G×E effects.

In chapter 8, G×E was estimated for two traits together, by investigating the dependency of the genetic relationship between milk yield and health and fertility on herd environment. Herd environment was described by 4 principal components comprising a number of herd characteristics: 1) intensity defined as average production per cow, 2) average fertility, 3)

148

Summary

farm size, and 4) relative performance indicating whether herds had good (poor) health and fertility despite a high (low) production. Data was analysed with a multi-trait reaction norm model. Genetic correlations of milk yield with fertility (days to first service and number of inseminations per service period), and somatic cell score, were across all herd environments antagonistic but the magnitude varied across environments. Genetic correlation of milk yield with DFS varied from 0.30 in small herds to 0.48 in herds with low average fertility. Correlations of yield with NINS varied from 0.18 in large herds to 0.64 in high fertility herds, and with SCS from 0.25 in herds with a high fertility relative to production to 0.47 in herds with a relatively low fertility. Selection in environments of average value resulted in different predicted responses across environments. For instance, selection for a decrease of NINS of 0.1 in an average production environment decreased milk yield by 35 kg in low but by 178 kg in high production herds.

Results of previous chapters are discussed together in chapter 9, with main emphasis on use of reaction norm methodology, influence of selection on robustness of dairy cows, implications of estimated G×E effects, and animal health risks associated with high milk yields. Reaction norm models were found to be able to deal with heterogeneous genetic variances, as well as genetic correlations across environments smaller than 1.0, and therefore enable to accurately predict breeding values and performance across environments. Random regression models and structured antedependence models were found to be both able to estimate G×E, although they both underestimated the genetic correlation between a trait expressed in different environments. It was discussed that a reasonable proportion of estimated G×E effects were related to herd effects that are somewhat related to nutrition or energy status of the cow. The herd effects body condition score, fat-to-protein ratio, and change in fat percentage during lactation are all associated with differences in energy status of cows. Further, it was discussed that selection for increased yield likely leads to increased environmental sensitivity for yield, whereas selection for increased fertility likely leads to decreased environmental sensitivity for fertility. This indicates that selection for fertility likely results in increased robustness with respect to fertility. Based on the estimated genetic parameters, within the range of considered herd environments one breeding program for yield and fertility traits seems sufficient, whereas for SCC and survival different breeding programs might be more effective. Finally, the results of this thesis were discussed in the light of other results from the research project "Animal health risks associated with high milk yield". During discussion meetings, most dairy farmers indicated that management and breeding should go hand in hand, while a reasonable proportion of the farmers indicated that negative side effects for selection can be counteracted by improving management. Based on the results in this thesis it was demonstrated, however, that on average this is not the case.

Samenvatting

Samenvatting

Het doel van dit proefschrift is het schatten van genotype-milieu interactie voor productie, vruchtbaarheids- en gezondheidskenmerken bij melkvee. Genotype-milieu interactie is het verschijnsel dat veranderingen in het management (voeding, huisvesting, klimaat, etc.) op een melkveebedrijf een verschillende uitwerking hebben op de prestaties van verschillende dieren. Onderlinge verschillen in prestaties kunnen groter of kleiner worden, en in sommige gevallen kan het zo zijn dat het ene dier beter presteert op het ene bedrijf, terwijl een ander dier beter presteert op een ander bedrijf. Dit kan leiden tot twee effecten waarmee in de melkveehouderijpraktijk rekening moet worden gehouden: 1) een verandering in management heeft niet hetzelfde effect op de prestaties van verschillende dieren, en 2) voor verschillende typen bedrijven zijn wellicht verschillende dieren gewenst. Het schatten van genotype-milieu interactie kan helpen bij het beantwoorden van belangrijke vragen rond robuustheid en gezondheidsrisico's van hoge melkproducties.

Onderzoek naar genotype-milieu interactie is tot nu toe voornamelijk gericht op productiekenmerken. Door het toenemende belang van gezondheids- en vruchtbaarheidskenmerken is de aandacht voor genotype-milieu interactie toegenomen. Om de prestaties van dochters van een stier op verschillende bedrijven goed te kunnen voorspellen, is het bijvoorbeeld belangrijk om te weten of mastitisgevoeligheid van dochters van een stier hetzelfde is in verschillende landen, of op bedrijven met een hoog en laag celgetal. Daarnaast is het belangrijk om te weten of selectie onder alle omstandigheden even betrouwbaar is. In dit proefschrift zijn de fokwaarden van stieren berekend met een zogenaamd reactienorm model waarmee de fokwaarden als functie van de omgeving zijn gemodelleerd. Het verloop van de fokwaarde van een stier over verschillende omgevingen heen, geeft aan of de dochters van een stier al dan niet gevoelig zijn voor veranderingen in de omgeving. In dit proefschrift zijn de reactienorm modellen getest met behulp van computer simulatie en vergeleken met een ander type model, het SAD model. Bij de analyse van praktijkdata zijn de reactienorm modellen uitgebreid van een kenmerk gecombineerd met een dimensie voor de omgeving, naar meerdere dimensies voor omgevingen en meerdere kenmerken. In dit laatste geval wordt er rekening mee gehouden dat de genetische correlatie tussen melkproductie en bijvoorbeeld vruchtbaarheid beïnvloedt wordt door de omgeving.

Verschillen in bedrijfsomgevingen zijn bepaald aan de hand van informatie die beschikbaar was via melkcontrole-, inseminatie- en exterieurgegevens van het NRS. Bedrijfsomgevingen zijn in dit proefschrift gedefinieerd als het gemiddelde van dieren op hetzelfde bedrijf voor een bepaald kenmerk, bijvoorbeeld bedrijfsjaargemiddelde eiwitproductie, tussenkalftijd, celgetal of het aantal melkgevende dieren. In eerste instantie is gekeken naar de relaties tussen deze bedrijfskenmerken en onderlinge verschillen in productie van vaarzen. Over het algemeen was de rangschikking van stieren, op basis van de prestaties van hun dochters, hetzelfde voor verschillende bedrijven. De onderlinge verschillen in fokwaarde voor productie tussen stieren waren echter wel groter op bedrijven met een

gemiddeld hoge productie, een hoge persistentie, een jonge leeftijd bij afkalven, een hoge conditiescore, een korte tussenkalftijd, en wanneer de meeste dieren in de winter afkalfden.

Vervolgens is gekeken naar de relaties tussen dezelfde bedrijfskenmerken en onderlinge verschillen in fokwaarden voor vruchtbaarheid, celgetal, celgetalpieken (voorspeller voor mastitis), survival (levensduur) en conditiescore. Bedrijven met gemiddeld een lage tussenkalftijd en een laag celgetal, scoorden het beste voor de genoemde gezondheids- en vruchtbaarheidskenmerken. Op bedrijven met een bovengemiddelde productie, scoorden de dieren bovengemiddeld voor de gezondheidskenmerken, maar tegelijkertijd was de vruchtbaarheid lager dan gemiddeld. De onderlinge verschillen in fokwaarden van stieren voor vruchtbaarheid namen fors toe, naarmate de gemiddelde vruchtbaarheid op een bedrijf slechter werd. Het bleek dat het effect van selectie voor vruchtbaarheid, t.o.v. eiwitproductie, twee keer zo groot was op bedrijven met een gemiddeld lage vruchtbaarheid, als op bedrijven met een gemiddeld goede vruchtbaarheid. De stieren die het beste scoorden voor vruchtbaarheid in een omgeving, scoorden doorgaans ook het beste in andere omgevingen. Voor celgetal, celgetalpieken en survival, was de rangschikking van stieren in een aantal situaties afhankelijk van de bedrijfsomgeving.

Celgetal is een kenmerk dat sterk afhangt van de omgeving (aanwezigheid van mastitisverwekkers, hygiëne tijdens het melken, etc.), niet alleen per lactatie, maar ook tijdens de lactatie. Daarom is voor celgetal ook bepaald hoe groot de genotype-milieu interactie is op testdagniveau. Celgetal hing hierbij af van het berekende tankcelgetal op de bedrijven. De resultaten lieten zien dat de rangschikking van stieren op basis van celgetal aan het begin van de lactatie werd beïnvloedt door tankcelgetal, terwijl dit aan het einde van de lactatie nauwelijks het geval was. De verschillen in fokwaarden van de stieren bleken groter op testdagniveau, dan op lactatieniveau. Dit geeft aan dat ook veranderingen in de omgeving op korte termijn (bijvoorbeeld de aanwezige mastitisverwekkers) kunnen leiden tot verschillen in expressie van de genetische aanleg voor celgetal.

Analoog aan de analyses voor celgetal is gekeken naar het verloop van vet percentage en kg vet als functie van bedrijfsgemiddelde vetpercentage, voor Australisch melkvee. Dit verloop is gebruikt als voorspeller voor de metabolische stoornis 'melkvet-onderdrukking', die mogelijk kan leiden tot (subklinische) pensverzuring en klauwbevangenheid. Hierbij is er vanuit gegaan dat het gemiddelde vetpercentage op een bedrijf samenhing met de samenstelling van het rantsoen: een laag vetpercentage werd geassocieerd met een hoog aandeel krachtvoer en weinig structuur in het rantsoen. Kenmerken gedefinieerd op basis van het verloop van vetpercentage, hadden een erfelijkheidsgraad van 4 tot 5% en gaven daarmee aan dat selectie tegen melkvetonderdrukking mogelijk is.

Uiteindelijk is ook de invloed van veranderingen in de omgeving op de relatie tussen verschillende kenmerken onderzocht. In alle omgevingen waren hoge fokwaarden voor melk gemiddeld geassocieerd met lagere vruchtbaarheid (hogere fokwaarde voor aantal

inseminaties) en hoger celgetal. De sterkte van de relatie tussen de fokwaarde voor melk en vruchtbaarheid en celgetal hing echter wel af van de bedrijfsomgeving. De relatie tussen fokwaarden voor melk en vruchtbaarheid was bijvoorbeeld zwak op grote bedrijven, maar sterk op bedrijven met gemiddeld een goede vruchtbaarheid.

De resultaten in dit proefschrift kunnen gebruikt worden om actuele vragen te beantwoorden over robuustheid van melkkoeien en de gezondheidsrisico's van hoge melkproducties. Het beeld bestaat dat fokken op een hogere productie ten koste is gegaan van robuustheid van de dieren. De resultaten in dit proefschrift laten zien dat de productie van dieren met een hoge aanleg voor melkproductie relatief gezien het sterkste verandert ten gevolge van veranderingen in de bedrijfsomgeving. Echter, de dieren met de hoogste aanleg voor melkproductie hebben, ondanks deze grotere veranderingen, wel de hoogste productie onder verschillende omstandigheden. De resultaten lieten ook zien, dat de vruchtbaarheid van dieren met een hoge aanleg voor vruchtbaarheid relatief gezien het minst verandert ten gevolge van veranderingen in de omgeving. Dat duidt erop dat fokken op vruchtbaarheid de robuustheid van dieren m.b.t. vruchtbaarheid verhoogd.

Om maatschappelijke zorgen over gezondheidsrisico's van hoogproducerende dieren te beantwoorden is in de algemene discussie de relatie gelegd tussen de geschatte genotype-milieu interactie in dit proefschrift en de overige resultaten van een breder project, dat bestond uit een dierexperiment, genetisch-epidemiologisch onderzoek en discussies met verschillende stakeholders. Tijdens discussiebijeenkomsten met veehouders bleek dat de meeste melkveehouders vinden dat management en fokkerij hand in hand moet gaan, terwijl sommigen denken dat negatieve effecten op vruchtbaarheid en gezondheid tengevolge van eenzijdig fokken op productie, kunnen worden opgevangen met managementveranderingen. Uit het onderzoek bleek dat managementveranderingen de verminderde vruchtbaarheid en verslechterde uiergezondheid als gevolg van selectie op productie niet kunnen opvangen en dat fokken op deze kenmerken daarom noodzakelijk is.

REFERENCES

Albers, G. A. A., K. Broekman, J. Visscher, N. Buddiger, and R. Maatman. 2002. Role of genotype x environment interactions in applied breeding programmes. 7th World Congress on Genetics Applied to Livestock Production, August 19-23. Montpellier, France.

Ali, A. K. A. and G. E. Shook. 1980. An optimum transformation for somatic cell concentration in milk. J. Dairy Sci. 63:487-490.

Banos, G. and G. E. Shook. 1990. Genotype by environment interaction and genetic correlations among parities for somatic-cell count and milk-yield. J. Dairy Sci. 73:2563-2573.

Bargo, F., L. D. Muller, J. E. Delahoy, and T. W. Cassidy. 2002. Milk response to concentrate supplementation of high producing dairy cows grazing at two pasture allowances. J. Dairy Sci. 85:1777-1792.

Bargo, F., L. D. Muller, E. S. Kolver, and J. E. Delahoy. 2003. Invited review: production and digestion of supplemented dairy cows on pasture. J. Dairy Sci. 86:1-42.

Barkema, H. W., Y. H. Schukken, T. Lam, M. L. Beiboer, G. Benedictus, and A. Brand. 1998a. Management practices associated with low, medium, and high somatic cell counts in bulk milk. J. Dairy Sci. 81:1917-1927.

Barkema, H. W., Y. H. Schukken, T. Lam, M. L. Beiboer, G. Benedictus, and A. Brand. 1999a. Management practices associated with the incidence rate of clinical mastitis. J. Dairy Sci. 82:1643-1654.

Barkema, H. W., Y. H. Schukken, T. Lam, M. L. Beiboer, H. Wilmink, G. Benedictus, and A. Brand. 1998b. Incidence of clinical mastitis in dairy herds grouped in three categories by bulk milk somatic cell counts. J. Dairy Sci. 81:411-419.

Barkema, H. W., J. D. Van der Ploeg, Y. H. Schukken, T. J. G. M. Lam, G. Benedictus, and A. Brand. 1999b. Management style and its association with bulk milk somatic cell count and incidence rate of clinical mastitis. J. Dairy Sci. 82:1655-1663.

Beerda, B., W. Ouweltjes, J. J. Windig, M. P. L. Calus, and R. F. Veerkamp. 2005. Dairy cow health and the effect of genetic merit for milk production, management and interactions between these: blood metabolites and enzymes. 56th Annual Meeting of the EAAP, June 5-8. Uppsala, Sweden.

Beilharz, R. G., B. G. Luxford, and J. L. Wilkinson. 1993. Quantitative genetics and evolution - is our understanding of genetics sufficient to explain evolution. J. Anim. Breed. Genet. 110:161-170.

Bennett, R. M. 1997. Farm animal welfare and food policy. Food Policy. 22:281-288.

Berglund, B. and B. Danell. 1987. Live weight changes, feed consumption, milk yield and energy balance in dairy cattle during the first period of lactation. Acta Agric. Scand. Sect. A-Anim. Sci. 37:195-509.

References

Berry, D. P., F. Buckley, P. Dillon, R. D. Evans, M. Rath, and R. F. Veerkamp. 2002. Genotype x environment interaction for body condition score, body weight and milk yield using random regression models. 7th World Congress on Genetics Applied to Livestock Production, August 19-23. Montpellier, France.

Berry, D. P., F. Buckley, P. Dillon, R. D. Evans, M. Rath, and R. F. Veerkamp. 2003. Estimation of genotype X environment interactions, in a grassbased system, for milk yield, body condition score, and body weight using random regression models. Livest. Prod. Sci. 83:191-203.

Boettcher, P. J., J. Fatehi, and M. M. Schutz. 2003. Genotype x environment interactions in conventional versus pasture-based dairies in Canada. J. Dairy Sci. 86:383-389.

Boldman, K. G. and A. E. Freeman. 1990. Adjustment for heterogeneity of variances by herd production level in dairy cow and sire evaluation. J. Dairy Sci. 73:503-512.

Bowman, P. J., P. M. Visscher, and M. E. Goddard. 1996. Customized selection indices for dairy bulls in Australia. Anim. Sci. 62:393-403.

Brotherstone, S. and W. G. Hill. 1986. Heterogeneity of variance amongst herds for milk production. Anim. Prod. 42:297-303.

Buckley, F., P. Dillon, M. Rath, and R. F. Veerkamp. 2000. The relationship between genetic merit for yield and live weight, condition score, and energy balance of spring calving Holstein Friesian dairy cows on grass based systems of milk production. J. Dairy Sci. 83:1878-1886.

Buenger, A., V. Ducrocq, and H. H. Swalve. 2001. Analysis of survival in dairy cows with supplementary data on type scores and housing systems from a region of northwest Germany. J. Dairy Sci. 84:1531-1541.

Calus, M. P. L., P. Bijma, and R. F. Veerkamp. 2004. Effects of data structure on the estimation of covariance functions to describe genotype by environment interactions in a reaction norm model. Genet. Sel. Evol. 36:489-507.

Calus, M. P. L., A. F. Groen, and G. de Jong. 2002. Genotype x environment interaction for protein yield in Dutch dairy cattle as quantified by different models. J. Dairy Sci. 85:3115-3123.

Calus, M. P. L., F. Jaffrezic, and R. F. Veerkamp. 2005a. Use of structured antedependence models to estimate genotype by environment interaction. 56th Annual Meeting of the EAAP, June 5-8. Uppsala, Sweden.

Calus, M. P. L., L. L. G. Janss, J. J. Windig, B. Beerda, and R. F. Veerkamp. 2005b. Effectiveness of selection for lower somatic cell count (SCC) in herds with different levels of SCC. 56th Annual Meeting of the EAAP, June 5-8. Uppsala, Sweden.

Calus, M. P. L. and R. F. Veerkamp. 2003. Estimation of environmental sensitivity of genetic merit for milk production traits using a random regression model. J. Dairy Sci. 86:3756-3764.

References

Calus, M. P. L., J. J. Windig, and R. F. Veerkamp. 2005c. Associations among descriptors of herd management and phenotypic and genetic levels of health and fertility. J. Dairy Sci. 88:2178-2189.

Carabaño, M. J., K. M. Wade, and L. D. Vanvleck. 1990. Genotype by environment interactions for milk and fat production across regions of the United-States. J. Dairy Sci. 73:173-180.

Carlén, E., K. Jansson, and E. Strandberg. 2005. Genotype by environment interaction for udder health traits in Swedish Holstein cattle. 56th Annual Meeting of the EAAP, June 5-8. Uppsala, Sweden.

Castillo-Juarez, H., P. A. Oltenacu, R. W. Blake, C. E. McCulloch, and E. G. Cienfuegos-Rivas. 2000. Effect of herd environment on the genetic and phenotypic relationships among milk yield, conception rate, and somatic cell score in Holstein cattle. J. Dairy Sci. 83:807-814.

Chilliard, Y., M. Cisse, R. Lefaivre, and B. Remond. 1991. Body-composition of dairy-cows according to lactation stage, somatotropin treatment, and concentrate supplementation. J. Dairy Sci. 74:3103-3116.

Cienfuegos-Rivas, E. G., P. A. Oltenacu, R. W. Blake, S. J. Schwager, H. Castillo-Juarez, and F. J. Ruiz. 1999. Interaction between milk yield of Holstein cows in Mexico and the United States. J. Dairy Sci. 82:2218-2223.

Costa, C. N., R. W. Blake, E. J. Pollak, P. A. Oltenacu, R. L. Quaas, and S. R. Searle. 2000. Genetic analysis of Holstein cattle populations in Brazil and the United States. J. Dairy Sci. 83:2963-2974.

Cromie, A. R. 1999. Genotype by environment interaction for milk production traits in Holstein Friesian dairy cattle in Ireland. PhD Thesis, Queens University of Belfast, Belfast, Ireland.

Crump, R. E., N. R. Wray, R. Thompson, and G. Simm. 1997. Assigning pedigree beef performance records to contemporary groups taking account of within-herd calving patterns. Anim. Sci. 65:193-198.

Dalley, D. 2002. Overcoming decreased milk solids production for Gippsland dairy herds in late winter/spring. Natural Resources and Environment, Agriculture Victoria Ellinbank, Australia.

De Haas, Y., H. W. Barkema, Y. H. Schukken, and R. F. Veerkamp. 2003. Genetic associations for pathogen-specific clinical mastitis and patterns of peaks in somatic cell count. Anim. Sci. 77:187-195.

De Haas, Y., H. W. Barkema, and R. F. Veerkamp. 2002. The effect of pathogen-specific clinical mastitis on the lactation curve for somatic cell count. J. Dairy Sci. 85:1314-1323.

References

De Haas, Y., R. F. Veerkamp, H. W. Barkema, Y. T. Gröhn, and Y. H. Schukken. 2004. Associations between pathogen-specific cases of clinical mastitis and somatic cell count patterns. J. Dairy Sci. 87:95-105.

De Veer, J. C. and L. D. Van Vleck. 1987. Genetic parameters for first lactation milk yields at three levels of herd production. J. Dairy Sci. 70:1434-1441.

De Vries, M. J. and R. F. Veerkamp. 2000. Energy balance of dairy cattle in relation to milk production variables and fertility. J. Dairy Sci. 83:62-69.

Dekkers, J. C. M., T. VanErp, and Y. H. Schukken. 1996. Economic benefits of reducing somatic cell count under the milk quality program of Ontario. J. Dairy Sci. 79:396-401.

Detilleux, J. and P. L. Leroy. 2000. Application of a mixed normal mixture model for the estimation of mastitis-related parameters. J. Dairy Sci. 83:2341-2349.

Dickerson, G. E. 1962. Implications of genetic environmental interaction in animal breeding. Anim. Prod. 4:47-64.

Dillon, P., F. Buckley, P. O'Connor, D. Hegarty, and M. Rath. 2003a. A comparison of different dairy cow breeds on a seasonal grass- based system of milk production 1. Milk production, live weight, body condition score and DM intake. Livest. Prod. Sci. 83:21-33.

Dillon, P., S. Snijders, F. Buckley, B. Harris, P. O'Connor, and J. F. Mee. 2003b. A comparison of different dairy cow breeds on a seasonal grass- based system of milk production 2. Reproduction and survival. Livest. Prod. Sci. 83:35-42.

Distl, O. 2001. Implications of health traits in breeding of dairy cattle. Arch. Tierz. 44:365-380.

Dockes, A. C. and F. Kling-Eveillard. 2004. Breeders representations of animals and animal welfare. 55th Annual Meeting of the EAAP, September 5-8. Bled, Slovenia.

Dodenhoff, J. and H. H. Swalve. 1998. Heterogeneity of variances across regions of northern Germany and adjustment in genetic evaluation. Livest. Prod. Sci. 53:225-236.

Dong, M. C. and I. L. Mao. 1990. Heterogeneity of (co)variance and heritability in different levels of intraherd milk production variance and of herd average. J. Dairy Sci. 73:843-851.

Emanuelson, U. 1988. Recording of production diseases in cattle and possibilities for genetic improvements - a review. Livest. Prod. Sci. 20:89-106.

Enevoldsen, C., J. Hindhede, and T. Kristensen. 1996. Dairy herd management types assessed from indicators of health, reproduction, replacement, and milk production. J. Dairy Sci. 79:1221-1236.

Erskine, R. J., R. J. Eberhart, L. J. Hutchinson, S. B. Spencer, and M. A. Campbell. 1988. Incidence and types of clinical mastitis in dairy herds with high and low somatic-cell counts. JAVMA. 192:761-765.

Esslemont, R. J. 2003. The costs of poor fertility and what to do about reducing them. Cattle Practice. 11:237-250.

References

Esslemont, R. J., M. A. Kossaibati, and J. Allcock. 2001. Economics of fertility in dairy cows. Pages 19-29 *in* Fertility in the High-Producing Dairy Cow. Br. Soc. Anim. Sci. Occas. Publ. 26.

Falconer, D. S. 1952. The problem of environment and selection. Am. Nat. 86:293-298.

Falconer, D. S. 1990. Selection in different environments - effects on environmental sensitivity (reaction norm) and on mean performance. Genet. Res. 56:57-70.

Falconer, D. S. and T. F. C. Mackay. 1996. Introduction to Quantitative Genetics. 4th ed. Longman Group, Essex, UK.

FAO. 1998. Secondary Guidelines for Development of National Farm Animal Genetic Resources Management Plans. Management of Small Populations at Risk. FAO, Rome.

Fatehi, J., S. Stella, J. J. Shannon, and P. J. Boettcher. 2003. Genetic parameters for feet and leg traits evaluated in different environments. J. Dairy Sci. 86:661-666.

Faye, B., F. Lescourret, N. Dorr, E. Tillard, B. MacDermott, and J. McDermott. 1997. Interrelationships between herd management practices and udder health status using canonical correspondence analysis. Prev. Vet. Med. 32:171-192.

Fikse, W. F. 2002. Advances in international genetic evaluation procedures of dairy cattle. PhD Thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Fikse, W. F., R. Rekaya, and K. A. Weigel. 2003a. Assessment of environmental descriptors for studying genotype by environment interaction. Livest. Prod. Sci. 82:223-231.

Fikse, W. F., R. Rekaya, and K. A. Weigel. 2003b. Genotype x environment interaction for milk production in Guernsey cattle. J. Dairy Sci. 86:1821-1827.

Foulley, J. L., B. Bouix, B. Goffinet, and J. M. Elsen. 1990. Connectedness in genetic evaluation. Pages 277-308 *in* Advances in statistical methods for genetic improvement of livestock. D. Gianola and K. Hammond, eds. Springer-Verlag, Berlin.

Gaynor, P. J., R. A. Erdman, B. B. Teter, J. Sampugna, A. V. Capuco, D. R. Waldo, and M. Hamosh. 1994. Milk fat yield and composition during abomasal infusion of cis or trans octadecenoates in Holstein cows. J. Dairy Sci. 77:157-165.

Gilmour, A. R., B. R. Cullis, S. J. Welham, and R. Thompson. 2002a. ASREML Reference Manual. New South Wales Agriculture, Orange Agricultural Institute, Orange, NSW, Australia.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, S. J. Welham, and R. Thompson. 2002b. ASReml User Guide Release 1.0. VSN International Ltd, Hemel Hempstead, UK.

Goddard, M. E. 1998. Consensus and debate in the definition of breeding objectives. J. Dairy Sci. 81:6-18.

Grieve, D. G., S. Korver, Y. S. Rijpkema, and G. Hof. 1986. Relationship between milk composition and some nutritional parameters in early lactation. Livest. Prod. Sci. 14:239-254.

References

Griinari, J. M., D. A. Dwyer, M. A. McGuire, D. E. Bauman, D. L. Palmquist, and K. V. Nurmela. 1998. Trans-octadecenoic acids and milk fat depression in lactating dairy cows. J. Dairy Sci. 81:1251-1261.

Groen, A. F., K. de Groot, J. D. van der Ploeg, and D. Roep. 1993. Stijlvol fokken: een oriënterende studie naar de relatie tussen sociaal-economische verscheidenheid en bedrijfsspecifieke fokdoeldefinitie. Vakgroep Veefokkerij en Vakgroep Rurale Sociologie, Landbouwuniversiteit Wageningen, Wageningen, The Netherlands.

Gröhn, Y. T. and M. L. Bruss. 1990. Effect of diseases, production, and season on traumatic reticuloperitonitis and ruminal acidosis in dairy cattle. J. Dairy Sci. 73:2355-2363.

Gröhn, Y. T., H. N. Erb, C. E. McCulloch, and H. S. Saloniemi. 1989. Epidemiology of metabolic disorders in dairy cattle: association among host characteristics, disease, and production. J. Dairy Sci. 72:1876-1885.

Haile-Mariam, M., M. E. Goddard, and P. J. Bowman. 2001. Estimates of genetic parameters for daily somatic cell count of Australian dairy cattle. J. Dairy Sci. 84:1255-1264.

Hayes, B. J., M. Carrick, P. Bowman, and M. E. Goddard. 2003. Genotype x environment interaction for milk production of daughters of Australian dairy sires from test-day records. J. Dairy Sci. 86:3736-3744.

Henderson, C. R. 1973. Sire evaluation and genetic trends. Anim. Breeding Genet. Symp. in Honor of Dr. J. L. Lush, Champaign, IL. 10.

Hill, W. G., M. R. Edwards, M. K. A. Ahmed, and R. Thompson. 1983. Heritability of milk yield and composition at different levels and variability of production. Anim. Prod. 36:59-69.

Hill, W. G. and R. Thompson. 1978. Probabilities of non-positive definite between-group or genetic covariance matrices. Biometrics. 34:429-439.

Ibáñez, M. A., M. J. Carabaño, and R. Alenda. 1999. Identification of sources of heterogeneous residual and genetic variances in milk yield data from the Spanish Holstein-Friesian population and impact on genetic evaluation. Livest. Prod. Sci. 59:33-49.

Ingvartsen, K. L., R. J. Dewhurst, and N. C. Friggens. 2003. On the relationship between lactational performance and health: is it yield or metabolic imbalance that cause production diseases in dairy cattle? A position paper*1. Livest. Prod. Sci. 83:277-308.

Jaffrezic, F. and S. D. Pletcher. 2000. Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. Genetics. 156:913-922.

Jaffrezic, F., E. Venot, D. Laloë, A. Vinet, and G. Renand. 2004. Use of structured antedependence models for the genetic analysis of growth curves. J. Anim. Sci. 82:3465-3473.

Jones, W. P., L. B. Hansen, and H. Chesterjones. 1994. Response of health-care to selection for milk-yield of dairy- cattle. J. Dairy Sci. 77:3137-3152.

References

Keady, T. W. J., C. S. Mayne, D. A. Fitzpatrick, and M. A. McCoy. 2001. Effect of concentrate feed level in late gestation on subsequent milk yield, milk composition, and fertility of dairy cows. J. Dairy Sci. 84:1468-1479.

Kellaway, R. C. and P. J. Colditz. 1975. The effect of heat stress on growth and nitrogen metabolism in Friesian and F1 Brahman x Friesian heifers. Aust. J. Agric. Res. 26:615-622.

Kennelly, J. J., B. Robinson, and G. R. Khorasani. 1999. Influence of carbohydrate source and buffer on rumen fermentation characteristics, milk yield, and milk composition in early-lactation Holstein cows. J. Dairy Sci. 82:2486-2496.

Kirkpatrick, M. and T. Bataillon. 1999. Artificial selection on phenotypically plastic traits. Gen. Res. 74:265-270.

Kirkpatrick, M. and N. Heckman. 1989. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. J. Math. Biol. 27:429-450.

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

Koivula, M., E. Negussie, and E. A. Mantysaari. 2004. Genetic parameters for test-day somatic cell count at different lactation stages of Finnish dairy cattle. Livest. Prod. Sci. 90:145-157.

Kolmodin, R., E. Strandberg, B. Danell, and H. Jorjani. 2004. Reaction norms for protein yield and days open in Swedish red and white dairy cattle in relation to various environmental variables. Acta Agric. Scand. Sect. A-Anim. Sci. 54:139-151.

Kolmodin, R., E. Strandberg, H. Jorjani, and B. Danell. 2003. Selection in the presence of a genotype by environment interaction : response in environmental sensitivity. Anim. Sci. 76:375-385.

Kolmodin, R., E. Strandberg, P. Madsen, J. Jensen, and H. Jorjani. 2002. Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. Acta Agric. Scand. Sect. A-Anim. Sci. 52:11-24.

Kolver, E. S. 2003. Nutritional limitations to increased production on pasture-based systems. Proc. Nutr. Soc. 62:291-300.

Kolver, E. S., J. R. Roche, M. J. de Veth, P. L. Thorne, and A. R. Napper. 2002. Total mixed rations versus pasture diets: evidence for a genotype x diet interaction in dairy cow performance. Proc. New Zealand Soc. Anim. Prod. 62:246-251.

Loeffler, S. H., M. J. de Vries, and Y. H. Schukken. 1999. The effects of time of disease occurrence, milk yield, and body condition on fertility of dairy cows. J. Dairy Sci. 82:2589-2604.

Lopez, H., L. D. Satter, and M. C. Wiltbank. 2004. Relationship between level of milk production and estrous behavior of lactating dairy cows. Anim. Reprod. Sci. 81:209-223.

References

Lowman, B. G., N. A. Scott, and S. H. Sommerville. 1976. Condition scoring of cattle - revised ed., East of Scotland Coll. of Agric, Edinburgh, Scotland.

Lynch, M. and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

Mallard, B. A., J. C. Dekkers, M. J. Ireland, K. E. Leslie, S. Sharif, C. L. Vankampen, L. Wagter, and B. N. Wilkie. 1998. Alteration in immune responsiveness during the peripartum period and its ramification on dairy cow and calf health. J. Dairy Sci. 81:585-595.

Mark, T. 2004. Applied genetic evaluations for production and functional traits in dairy cattle. J. Dairy Sci. 87:2641-2652.

Mathur, P. K. 2002. Methods for estimation and use of genotype-environment interactions. 7th World Congress on Genetics Applied to Livestock Production, August 19-23. Montpellier, France.

Mathur, P. K. and P. Horst. 1994. Genotype by environment interactions in laying hens based on relationship between breeding values of sires in temperate and tropical environments. Poult. Sci. 73:1777-1784.

Merks, J. W. M., E. W. Brascamp, and J. B. M. Wilmink. 1985. Genotype environment interaction in pig breeding programmes: methods of estimation and relevance of the estimates. Livest. Prod. Sci. 13:135-146.

Meuwissen, T. H. E. 1990. Optimization of dairy cattle breeding plans with increased female reproductive rates. PhD Thesis, Landbouwuniversiteit Wageningen, Wageningen, The Netherlands.

Meuwissen, T. H. E., G. De Jong, and B. Engel. 1996. Joint estimation of breeding values and heterogeneous variances of large data files. J. Dairy Sci. 79:310-316.

Meyer, K. 1987. Estimates of variances due to sire x herd interactions and environmental covariances between paternal half-sibs for first lactation dairy production. Livest. Prod. Sci. 17:95-115.

Miltenburg, J. D., D. deLange, A. P. P. Crauwels, J. H. Bongers, M. J. M. Tielen, Y. H. Schukken, and A. R. W. Elbers. 1996. Incidence of clinical mastitis in a random sample of dairy herds in the southern Netherlands. Vet. Rec. 139:204-207.

Mulder, H. A. and P. Bijma. 2005. Effects of genotype x environment interaction on genetic gain in breeding programs. J. Anim. Sci. 83:49-61.

Mulder, H. A., A. F. Groen, G. De Jong, and P. Bijma. 2004. Genotype x environment interaction for yield and somatic cell score with automatic and conventional milking systems. J. Dairy Sci. 87:1487-1495.

Namkoong, G. 1985. The influence of composite traits on genotype by environment relations. Theor. Appl. Genet. 70:315-317.

References

Nauta, W. J., T. Baars, and H. Bovenhuis. Converting to organic dairy farming: Consequences for production, somatic cell scores and calving interval of first parity Holstein cows. Livest. Prod. Sci. (in press).

Neave, F. K., F. H. Dodd, R. G. Kingwill, and D. R. Westgarth. 1969. Control of mastitis in dairy herds by hygiene and management. J. Dairy Sci. 52:696-707.

Nocek, J. E. 1997. Bovine acidosis: implications on laminitis. J. Dairy Sci. 80:1005-1028.

NRS. 2001a. Subject: E-17 Fokwaarde Vruchtbaarheid. https://www.cr-delta.nl/crd-nrs/index.htm. Accessed Sep. 14, 2004.

NRS. 2001b. Subject: E- 9 Index Netto Melkgeld (INET). https://www.cr-delta.nl/crd-nrs/index.htm. Accessed Sep. 14, 2004.

Nunez-Anton, V. and D. L. Zimmerman. 2000. Modeling nonstationary longitudinal data. Biometrics. 56:699-705.

O'Connell, J. M., F. Buckley, M. Rath, and P. Dillon. 2000. The effects of cow genetic merit and feeding treatment on milk production, herbage intake and grazing behaviour of dairy cows. Irish J. Agr. Food Res. 39:369-381.

Odegard, J., J. Jensen, G. Klemetsdal, P. Madsen, and B. Heringstad. 2003. Genetic analysis of somatic cell score in Norwegian cattle using random regression test-day models. J. Dairy Sci. 86:4103-4114.

Oseni, S., I. Misztal, S. Tsuruta, and R. Rekaya. 2004. Genetic components of days open under heat stress. J. Dairy Sci. 87:3022-3028.

Ouweltjes, W., B. Beerda, J. J. Windig, M. P. L. Calus, and R. F. Veerkamp. 2005. Dairy cow health and the effect of genetic merit for milk production, management and interactions between these: udder health parameters. 56th Annual Meeting of the EAAP, June 5-8. Uppsala, Sweden.

Petersson, K. J., R. Kolmodin, and E. Strandberg. 2005. Genotype by environment interaction for length of productive life in Swedish Red and White dairy cattle. Acta Agric. Scand. Sect. A-Anim. Sci. 55:9-15.

Philipsson, J., G. Banos, and T. Arnason. 1994. Present and future uses of selection index methodology in dairy cattle. J. Dairy Sci. 77:3252-3261.

Philipsson, J., G. Ral, and B. Berglund. 1995. Somatic-cell count as a selection criterion for mastitis resistance in dairy-cattle. Livest. Prod. Sci. 41:195-200.

Pollott, G. E. and J. C. Greeff. 2004. Genotype x environment interactions and genetic parameters for fecal egg count and production traits of Merino sheep. J. Anim. Sci. 82:2840-2851.

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

References

Pool, M. H., V. E. Olori, M. P. L. Calus, and R. F. Veerkamp. 2003. Aspects of milk yield adjustment in the parameter estimation for genetic evaluation of survival. Proc. Mtg. Int. Bull. Eval. Serv, March 2-3. Beltsville, MD, USA. Bulletin no 30:p 25-28.

Powell, R. L. and P. M. VanRaden. 2002. International dairy bull evaluations expressed on national, subglobal, and global scales. J. Dairy Sci. 85:1863-1868.

Productschap Zuivel. 2004. Subject: Ontwerp-Verordening tot wijziging (1) van de Zuivelverordening 2003, Vaststelling frequentie en beoordeling resultaten kwaliteitsonderzoek. http://www.prodzuivel.nl/index.asp?frame=http%3A//www.prodzuivel.nl/pz/productschap/bestuur/20030618/wijziging%25201%2520vaststelling.htm. Accessed May. 4, 2004.

Pryce, J. E., R. J. Esslemont, R. Thompson, R. F. Veerkamp, M. A. Kossaibati, and G. Simm. 1998. Estimation of genetic parameters using health, fertility and production data from a management recording system for dairy cattle. Anim. Sci. 66:577-584.

Pryce, J. E., B. L. Nielsen, R. F. Veerkamp, and G. Simm. 1999. Genotype and feeding system effects and interactions for health and fertility traits in dairy cattle. Livest. Prod. Sci. 57:193-201.

Pryce, J. E., M. D. Royal, P. C. Garnsworthy, and I. L. Mao. 2004. Fertility in the high-producing dairy cow. Livest. Prod. Sci. 86:125-135.

Pryce, J. E., R. F. Veerkamp, R. Thompson, W. G. Hill, and G. Simm. 1997. Genetic aspects of common health disorders and measures of fertility in Holstein Friesian dairy cattle. Anim. Sci. 65:353-360.

Raffrenato, E., R. W. Blake, P. A. Oltenacu, J. Carvalheira, and G. Licitra. 2003. Genotype by environment interaction for yield and somatic cell score with alternative environmental definitions. J. Dairy Sci. 86:2470-2479.

Rauw, W. M., E. Kanis, E. N. Noordhuizen-Stassen, and F. J. Grommers. 1998. Undesirable side effects of selection for high production efficiency in farm animals: a review. Livest. Prod. Sci. 56:15-33.

Ravagnolo, O. and I. Misztal. 2000. Genetic component of heat stress in dairy cattle, parameter estimation. J. Dairy Sci. 83:2126-2130.

Robertson, A. 1959. The sampling variance of the genetic correlation coefficient. Biometrics. 15:469-485.

Rougoor, C. W., W. J. A. Hanekamp, A. A. Dijkhuizen, M. Nielen, and J. B. M. Wilmink. 1999. Relationships between dairy cow mastitis and fertility management and farm performance. Prev. Vet. Med. 39:247-264.

Roxstrom, A., E. Strandberg, B. Berglund, U. Emanuelson, and J. Philipsson. 2001. Genetic and environmental correlations among female fertility traits and milk production in different parities of Swedish red and white dairy cattle. Acta Agric. Scand. Sect. A-Anim. Sci. 51:7-14.

References

Royal, M. D. and A. P. F. Flint. 2004. Genetic improvement of dairy cattle fertility. Cattle Practice. 12:21-29.

Samore, A. B., J. A. M. Van Arendonk, and A. F. Groen. 2001. Impact of area and sire by herd interaction on heritability estimates for somatic cell count in Italian Holstein Friesian cows. J. Dairy Sci. 84:2555-2559.

Sandoe, P. 2004. "Happy pigs are dirty" - conflicting perspectives on animal welfare. 55th Annual Meeting of the EAAP, September 5-8. Bled, Slovenia.

Sandoe, P., B. L. Nielsen, L. G. Christensen, and P. Sorensen. 1999. Staying good while playing god - The ethics of breeding farm animals. Anim. Welf. 8:313-328.

Sartori, R., R. Sartor-Bergfelt, S. A. Mertens, J. N. Guenther, J. J. Parrish, and M. C. Wiltbank. 2002. Fertilization and early embryonic development in heifers and lactating cows in summer and lactating and dry cows in winter. J. Dairy Sci. 85:2803-2812.

Schaeffer, L. R. 1994. Multiple-country comparison of dairy sires. J. Dairy Sci. 77:2671-2678.

Schaeffer, L. R. and J. C. M. Dekkers. 1994. Random regressions in animal models for test-day production in dairy cattle. Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, 7-12 August. Guelph. 18:443-446.

Schutz, M. M., P. M. Vanraden, and G. R. Wiggans. 1994. Genetic variation in lactation means of somatic cell scores for six breeds of dairy cattle. J. Dairy Sci. 77:284-293.

Simm, G., R. F. Veerkamp, and P. Persaud. 1994. The economic-performance of dairy-cows of different predicted genetic merit for milk solids production. Anim. Prod. 58:313-320.

Sölkner, J. and J. W. James. 1994. Curvilinearity in the relationship between traits competing for resources: a genetic model. Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, 7-12 August. Guelph, Canada. 19:151-154.

Sölkner, J., H. Nakimbugwe, and A. Valle Zarate. 1998. Analysis of determinants for succes and failure of village breeding programmes. Proceedings of the 6th World Congress on Genetics Applied to Livestock Production, 11-16 January. Armidale, Australia. 25:273-280.

Stearns, S., G. Dejong, and B. Newman. 1991. The effects of phenotypic plasticity on genetic correlations. Trends in Ecology & Evolution. 6:122-126.

Steinheim, G., T. Adnoy, T. Meuwissen, and G. Klemetsdal. 2004. Indications of breed by environment interaction for lamb weights in Norwegian sheep breeds. Acta Agric. Scand. Sect. A-Anim. Sci. 54:193-196.

Stockdale, C. R., A. Callaghan, and T. E. Trigg. 1987. Feeding high energy supplements to pasture-fed dairy cows. Effect of stage of lactation and level of supplement. Aust. J. Agric. Res. 38:927-940.

References

Strandberg, E., R. Kolmodin, P. Madsen, J. Jensen, and H. Jorjani. 2000. Genotype by environment interaction in Nordic dairy cattle studied by use of reaction norms. 2000 Interbull meeting, May 14-15. Bled Slovenia. 41-45.

Stratton, D. A. 1998. Reaction norm functions and QTL-environment interactions for flowering time in Arabidopsis thaliana. Heredity. 81:144-155.

Sutton, J. D. 1989. Altering milk composition by feeding. J. Dairy Sci. 72:2801-2814.

Tong, A. K. W., B. W. Kennedy, and J. E. Moxley. 1977. Sire by herd interaction for milk yield and composition traits. Can. J. Anim. Sci. 57:383.

Van der Waaij, E. H. 2004. A resource allocation model describing consequences of artificial selection under metabolic stress. J. Anim. Sci. 82:973-981.

Van der Waaij, E. H., L. L. G. Janss, and P. Bijma. 2004. Impact of selection environment on the evolution of environmental sensitivity. 55th Annual Meeting of the EAAP, September 5-8. Bled, Slovenia.

Van der Werf, J. H. J., M. E. Goddard, and K. Meyer. 1998. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. J. Dairy Sci. 81:3300-3308.

Van der Werf, J. H. J. and J. Ten Napel. 1991. Estimation of genotype-environment interaction for milk production under Dutch circumstances. 42nd Annual Meeting of the EAAP.

Van Tienderen, P. H. and H. P. Koelewijn. 1994. Selection on reaction norms, genetic correlations and constraints. Gen. Res. 64:115-125.

Veerkamp, R. F. 1998. Selection for economic efficiency of dairy cattle using information on live weight and feed intake: A review. J. Dairy Sci. 81:1109-1119.

Veerkamp, R. F. and M. E. Goddard. 1998. Covariance functions across herd production levels for test day records on milk, fat, and protein yields. J. Dairy Sci. 81:1690-1701.

Veerkamp, R. F., G. Simm, and J. D. Oldham. 1994. Effects of interaction between genotype and feeding system on milk-production, feed-intake, efficiency and body tissue mobilization in dairy-cows. Livest. Prod. Sci. 39:229-241.

Veerkamp, R. F., G. Simm, and J. D. Oldham. 1995. Genotype by environment interactions: experience from Langhill. Pages 59-66 *in* Breeding and Feeding the High Genetic Merit Dairy Cow. Br. Soc. Anim. Sci. Occas. Publ. 19.

Veerkamp, R. F., A. W. Stott, W. G. Hill, and S. Brotherstone. 1998. The economic value of somatic cell count payment schemes for UK dairy cattle breeding programmes. Anim. Sci. 66:293-298.

Visscher, P. M. and W. G. Hill. 1992. Heterogeneity of variance and dairy cattle breeding. Anim. Prod. 55:321-329.

Weigel, K. A. and T. J. Lawlor. 1994. Adjustment for heterogeneous variance in genetic evaluations for conformation of United-States Holsteins. J. Dairy Sci. 77:1691-1701.

166

References

Weigel, K. A. and R. Rekaya. 2000. A multiple-trait herd cluster model for international dairy sire evaluation. J. Dairy Sci. 83:815-821.

Weller, J. I., A. Saran, and Y. Zeliger. 1992. Genetic and environmental relationships among somatic-cell count, bacterial-infection, and clinical mastitis. J. Dairy Sci. 75:2532-2540.

Wicks, H. C. F. and J. D. Leaver. Influence of genetic merit and environment on somatic cell counts of Holstein-Friesian cows. Vet. J. (in press).

Wiggans, G. R. and P. M. VanRaden. 1991. Method and effect of adjustment for heterogeneous variance. J. Dairy Sci. 74:4350-4357.

Wilson, D. J., H. H. Das, R. N. Gonzalez, and P. M. Sears. 1997. Association between management practices, dairy herd characteristics, and somatic cell count of bulk tank milk. JAVMA. 210:1499-1502.

Windig, J. J., M. P. L. Calus, B. Beerda, W. Ouweltjes, and R. F. Veerkamp. 2005a. Environmental influences on genetic and phenotypic relationships between production and health and fertility in Dutch dairy cows. 56th Annual Meeting of the EAAP, June 5-8. Uppsala, Sweden.

Windig, J. J., M. P. L. Calus, G. de Jong, and R. F. Veerkamp. 2005b. The association between somatic cell count patterns and milk production prior to mastitis. Livest. Prod. Sci. 96:291-299.

Windig, J. J., M. P. L. Calus, and R. F. Veerkamp. 2005c. Influence of herd environment on health and fertility and their relationship with milk production. J. Dairy Sci. 88:335-347.

Woltereck, R. 1909. Weitere experimentelle Untersüchungen über Artveränderung, speziell über das Wesen quantitativer Artunterschiede bei Daphniden. Verhandlungen der Deutschen Zoologischen Gesellschaft. 19:110-172.

Zwald, N. R., K. A. Weigel, W. F. Fikse, and R. Rekaya. 2001. Characterization of dairy production systems in countries that participate in the International Bull Evaluation Service. J. Dairy Sci. 84:2530-2534.

Zwald, N. R., K. A. Weigel, W. F. Fikse, and R. Rekaya. 2003. Application of a multiple-trait herd cluster model for genetic evaluation of dairy sires from seventeen countries. J. Dairy Sci. 86:376-382.

ABBREVIATION KEY

AHTDF%	Average herd-test-day milk fat percentage
BCS	Body condition score
BMSCC	Bulk milk somatic cell count
BVM	Bivariate repeatability model
CF	Covariance functions
CIV	Calving interval
CV	Coefficient of variation
DFS	Days to first service
DFLS	Days first to last service
DIM	Days in milking
DLS	Days to last service
EBV	Estimated breeding value
EP	Environmental parameter
ES	Environmental sensitivity (of genetic variance)
FSC	First-service conception
G×E	Genotype × environment interaction
HYS	Herd-year-season
MF%	Milk fat percentage
MFD	Milk fat depression
MFDLAC	Milk fat depression trait defined on a lactation level
MFDTD	Milk fat depression trait defined on a test-day level
NINS	Number of inseminations per service period
NR56	Non-return at 56 d after first insemination
PCA	Principal components analysis
PC#	Principal component with number
pc$_{env\#}$	Average principal component value of environmental group number #
PTA	Predicted transmitting ability
RR	Random regression
RRM	Random regression model
SAD	Structured antedependence models
SCC	Somatic cell count
SCS	Somatic cell score
TBV	True breeding value

List of publications

Refereed

Schaeffer, L., M. Calus, en X. L. Liu. 2001. Genetic evaluation of conformation traits using random contemporary groups and reducing the influence of parent averages. Livestock Production Science 69: 129-137.

Calus, M. P. L., A. F. Groen, en G. de Jong. 2002. Genotype × environment interaction for protein yield in Dutch dairy cattle as quantified by different models. Journal of Dairy Science 85: 3115-3123.

Calus, M. P. L. en R. F. Veerkamp. 2003. Estimation of environmental sensitivity of genetic merit for milk production traits using a random regression model. Journal of Dairy Science 86: 3756-3764.

Calus, M. P. L., P. Bijma, en R. F. Veerkamp. 2004. Effects of data structure on the estimation of covariance functions to describe genotype by environment interactions in a reaction norm model. Genetics Selection Evolution 36: 489-507.

Calus, M. P. L., M. J. Carrick, R. F. Veerkamp, en M. E. Goddard. 2005. Estimation of genetic parameters for milk fat depression in dairy cattle. Journal of Dairy Science 88: 1166-1177.

Calus, M. P. L., J. J. Windig, en R. F. Veerkamp. 2005. Associations between descriptors of herd management and phenotypic and genetic levels of health and fertility. Journal of Dairy Science 88: 2178-2189.

Calus, M. P. L., L. L. G. Janss, en R. F. Veerkamp. Genotype by environment interaction of somatic cell count across bulk milk somatic cell count and days in milk. Submitted to Journal of Dairy Science.

Windig, J. J., M. P. L. Calus, en R. F. Veerkamp. 2005. Influence of herd environment on health and fertility and their relationship with milk production. Journal of Dairy Science 88: 335-347.

Windig, J. J., M. P. L. Calus, G. de Jong, en R. F. Veerkamp. 2005. The association between somatic cell count patterns and milk production prior to mastitis. Livestock Production Science 96: 291-299.

Windig, J. J., M. P. L. Calus, en R. F. Veerkamp. Genetic correlations between milk production and health and fertility dependent on herd environment. Journal of Dairy Science (in press).

Popular

Calus, M.P.L., en R.F. Veerkamp. 2003. Ieder bedrijf een eigen fokwaarde? Veeteelt, april 1, 2003. Pages 58-59.

Calus, M.P.L., en R.F. Veerkamp. 2004. Genetische selectie op pensverzuring. Veeteelt, September 2, 2004. Page 69.

Proceedings of international conferences and seminars

Calus, M.P.L., en R.F. Veerkamp. 2003. Environmental sensitivity of genetic merit for milk, fat and protein yield estimated by a random regression model. Proc. Joint Annual Meeting ADSA-ASAS, 22-26 June, 2003, Phoenix, Arizona, USA. pp. 37 (Communication 149).

Calus, M.P.L., J.J. Windig, en R.F. Veerkamp. 2004. Effect of herd environment on phenotypic and genetic levels of survival. Proc. 55th Annual meeting of the EAAP, 5-8 September, 2004, Bled, Slovenia. pp. 17 (Communication GM2.10).

Calus, M. P. L., L. L. G. Janss, J. J. Windig, B. Beerda, en R. F. Veerkamp. 2005. Effectiveness of selection for lower somatic cell count (SCC) in herds with different levels of SCC. Proc. 56th Annual Meeting of the EAAP, June 5-8, Uppsala, Sweden. pp. 343. (Communication G7.14).

Calus, M. P. L., F. Jaffrézic, en R.F. Veerkamp. 2005. Use of structured antedependence models to estimate genotype by environment interaction. Proc. 56th Annual Meeting of the EAAP, 5-8 June, Uppsala, Sweden. pp. 322. (Communication G5.10).

List of publications

Presentations at (international) conferences and seminars

Calus, M.P.L., and R.F. Veerkamp. 2003. Covariance functions are affected by non random use of sires. International Cattle Breeders Round Table, 15-17 December, Edinburgh, Scotland.

Calus, M.P.L., and R.F. Veerkamp. 2004. Nieuwe modellen voor het schatten van genotype-milieu interactie. F&G connection, 18-19 November, Vught, The Netherlands.

Calus, M.P.L., and R.F. Veerkamp. 2005. New models to estimate genotype by environment interaction. WIAS Science Day 2005, 17 February, Wageningen, The Netherlands.

Calus, M.P.L., Windig, J.J., and R.F. Veerkamp. 2006. New insights in G×E for yield, health and fertility from reaction norm models. 3rd International Cattle Breeders Round Table, 16-18 January, Milan, Italy.

Others

Calus, M.P.L. 1999. Rendabiliteit van de aanschaf van een krachtvoercomputer. MSc-thesis, Wageningen University, Wageningen, The Netherlands.

Calus, M.P.L. 2001. Genotype by Environment Interaction in Dutch Dairy Cattle studied by use of Reaction Norms. MSc-thesis, Wageningen University, Wageningen, The Netherlands.

CURRICULUM VITEA

Mario Pieter Lea Calus werd geboren op 1 mei 1978 in Oostburg (West Zeeuwsch-Vlaanderen). In 1996 behaalde hij zijn VWO-diploma aan het toenmalige 'Zwincollege' te Oostburg, waarna hij aan de universiteit in Wageningen Zoötechniek ging studeren. In 2000 deed hij 6 maanden stage in Canada, bij de Universiteit in Guelph (Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science). Tijdens de studie deed hij een klein afstudeervak bij de leerstoelgroep Agrarische Bedrijfseconomie en een groot afstudeervak bij de leerstoelgroep Fokkerij en Genetica. Na het behalen van het diploma, in september 2001, is hij zes maanden als junior onderzoeker werkzaam geweest bij de Divisie Dier en Omgeving bij ID-Lelystad (nu Animal Sciences Group). Vanaf april 2002 is hij werkzaam geweest als Assistent in Opleiding (AIO) bij de de leerstoelgroep Fokkerij en Genetica (onderzoekschool Wageningen Institute of Animal Science (WIAS)), gedetacheerd bij de Divisie Dier en Omgeving van ID-Lelystad (nu: Divisie Veehouderij van de Animal Sciences Group). Sinds 1 februari 2006 is hij werkzaam als post-doc onderzoeker bij het cluster Diergenetica en Biodiversiteit van de Divisie Veehouderij van de Animal Sciences Group in Lelystad.

NAWOORD

Na vier jaar werken is het proefschrift af. Een proefschrift schrijf je niet alleen. Het is misschien een cliché, maar het vat datgene wat ik hier wil zeggen wel mooi samen.

De interesse voor fokkerij was onvermijdelijk: enerzijds kon ik altijd al beter overweg met getalletjes dan met letters en anderzijds ontstond na het verhuizen naar de boerderij in 1988 steeds meer de interesse voor melkvee. De ultieme combinatie van die twee is de veefokkerij, dat was halverwege het VWO al wel duidelijk. Een ieder die jaren terug al bijgedragen heeft aan een van deze interesses, wil ik vanaf deze plaats hartelijk bedanken. Dat het wetenschappelijk onderzoek een voor de hand liggende keuze was, werd mijzelf pas veel later duidelijk.

Allereerst de leden van de begeleidingscommissie: Johan, Piter, Jack en Roel. Van harte bedankt voor alle opmerkingen, aanwijzingen en voor het helpen ophelderen van vele onduidelijkheden, al dan niet door mij gecreëerd. Roel, jou ben ik in het bijzonder zeer erkentelijk vanwege het haast altijd beschikbaar zijn ondanks alle drukke werkzaamheden en vanwege het vermogen om tegelijkertijd vol vertrouwen en continu kritisch te zijn. Verder dank ik de directe collega's en oud-collega's van het cluster Diergenetica en Biodiversiteit, en in het bijzonder alle kamergenoten die ik in deze vier jaar heb gehad, voor de collegialiteit en het helpen bij allerlei problemen.

Verder dank ik alle mensen uit het projectteam van "Gezondheidsrisico's hoge producties", en in het bijzonder Jack, Wijbrand, Bonne, en Roel voor de samenwerking tijdens de veehoudersbijeenkomsten. Thanks to all of the co-authors on my papers and conference proceedings contributing to this thesis: Piter Bijma, Mick Carrick, Mike Goddard, Florence Jaffrézic, Lucc Janss, Roel Veerkamp and Jack Windig.

Ten laatste dank aan familie en vrienden. Allereerst de studiegenoten uit Wageningen, maar ook de vrienden uit West Zeeuwsch-Vlaanderen, waarmee ter ontspanning menig uurtje in de juiste omgeving is doorgebracht, al dan niet hopend op een interactie. Zeer in het bijzonder ben ik dank verschuldigd aan mijn vader en moeder, en natuurlijk mijn zus Veronique, voor de nimmer aflatende steun voor mijn studie en werkzaamheden, ondanks de soms vele onduidelijkheden daaromheen.

Mario Calus

PhD Education plan

Training and Supervision Plan		
Name	Mario Calus	
Group	Animal Breeding and Genetics	
Daily supervisor	Roel Veerkamp	
Supervisors	Roel Veerkamp, Johan van Arendonk, Piter Bijma, Jack Windig	
eriod	April 2002 until Januari 2006	

The Graduate School
WAGENINGEN INSTITUTE of ANIMAL SCIENCES

The Basic Package (minimum 3 ECTS[1])	year	ECTS[1]
WIAS Introduction Course (mandatory)	2003	
WIAS course 'Philosophy of Science and Ethics' (mandatory)	2003	
Subtotal Basic Package		**3**

Scientific Exposure (conferences, seminars and presentations, minimum 8 ECTS)	year	ECTS
<u>International conferences (minimum 3 ECTS)</u>		
7th World Congress on Genetics Applied to Livestock Production, Montpellier, France	2002	
ADSA ASAS Joint Annual Meeting, Phoenix, Arizona	2003	
International Cattle Breeders Round Table, Scotland	2003	
Annual meeting of the EAAP, Bled, Slovenia	2004	
Annual meeting of the EAAP, Uppsala, Sweden	2005	
<u>Seminars and workshops</u>		
WIAS Science Day 2002, Wageningen, The Netherlands	2002	
Fokkerij&Genetica Connection, Vught, The Netherlands	2002	
PhD-Retreat "Gateway to the future"	2002	
Breeding for health in livestock (WIAS)	2003	
WIAS Science Day 2003, Wageningen, The Netherlands	2003	
Studiemiddag 'Koeien in Beweging. Nieuwe inzichten in klauwproblemen en locomotiestoornissen' Wageningen, The Netherlands	2003	
Scientific publishing: An introductory workshop for PhD students and young authors, WGS, Wageningen, The Netherlands	2004	
Fokkerij&Genetica Connection, Vught, The Netherlands	2004	
WIAS seminar "Farm Animal Genomics: from sequence to application"	2004	
WIAS Science Day 2005, Wageningen, The Netherlands	2005	
<u>Presentations (minimum 4 original presentations of which at least 1 oral, 1 ECTS each)</u>		
ADSA ASAS Joint Annual Meeting, Phoenix, Arizona (oral presentation)	2003	
International Cattle Breederes Round Table, Scotland (oral presentation)	2003	
Annual meeting of the EAAP, Bled, Slovenie (oral presentation)	2004	
Fokkerij&Genetica Connection 2004 (oral presentation)	2004	
WIAS Science Day 2005, Wageningen, The Netherlands (oral presentation)	2005	
Annual meeting of the EAAP, Uppsala, Sweden (poster)	2005	
Annual meeting of the EAAP, Uppsala, Sweden (oral presentation)	2005	
Subtotal International Exposure		**17**

PhD Education plan

In-Depth Studies (minimum 6 ECTS)	year	ECTS
Disciplinary and interdisciplinary courses		
Basic Statistical Methods for Longitudinal Data Analysis (Montpellier, France)	2002	
Quantitative Genetics in Animal Breeding (NOVA postgraduate course 2003)	2003	
WIAS Course "Genetic Algorithms applied to Animal Breeding"	2003	
Armidale Animal Breeding Summer Course (Practical Bayes for Beginners,		
Case Studies in QTL mapping using Bayesian Analysis, Advanced Genome analysis &		
Essential bioinformatics for animal geneticists)	2004	
Course "Theory and Practice of International Genetic Evaluation" (Interbull)	2005	
Course "Genes and environment" (Uppsala, Sweden)	2005	
WIAS Course "The biological basis for improved management and selection tools"	2005	
WIAS Course "QTL detection and fine mapping in complex pedigrees"	2005	
Subtotal In-Depth Studies		**16**
Professional Skills Support Courses (minimum 3 ECTS)	year	ECTS
WIAS Course Techniques for Scientific Writing	2002	
Course Writing for Academic Publication, Lelystad	2002	
WIAS midterm job assessment	2005	
NIBI cursus Onderzoeksmanagement	2002	
Subtotal Professional Skills Support Courses		**6**
Research Skills Training (apart from carrying out the PhD project, optional)	year	ECTS
Preparing own PhD research proposal (optional, maximum 6 ECTS)	2002	
External training period in Melbourne, Australia	2004	
Subtotal Research Skills Training		**8**
Education and Training Total		**50**

[1] one ECTS credit point equals a study load of approximately 28 hours.

NOTES