COMMUNAUTÉ FRANÇAISE DE BELGIQUE
UNIVERSITÉ DE LIÈGE – GEMBLOUX AGRO-BIO TECH

IMPROVEMENT OF GENETIC EVALUATION SYSTEMS FOR MATERNALLY INFLUENCED TRAITS AND MULTI-BREED LIVESTOCK POPULATIONS

SYLVIE VANDERICK

Essai présenté en vue de l'obtention du grade de
docteur en sciences agronomiques et ingénierie biologique

Promoteur: Nicolas Gengler

2017

COMMUNAUTÉ FRANÇAISE DE BELGIQUE
UNIVERSITÉ DE LIÈGE – GEMBLOUX AGRO-BIO TECH

IMPROVEMENT OF GENETIC EVALUATION SYSTEMS FOR MATERNALLY INFLUENCED TRAITS AND MULTI-BREED LIVESTOCK POPULATIONS

SYLVIE VANDERICK

Essai présenté en vue de l'obtention du grade de
docteur en sciences agronomiques et ingénierie biologique

Promoteur: Nicolas Gengler
2017

In loving memory of my dad,

Deep in my heart
You'll always stay.
Loved and missed
Everyday.

VANDERICK SYLVIE. (2017). IMPROVEMENT OF GENETIC EVALUATION SYSTEMS FOR MATERNALLY INFLUENCED TRAITS AND MULTI-BREED LIVESTOCK POPULATIONS. (PHD DISSERTATION IN ENGLISH). GEMBLOUX, BELGIUM, GEMBLOUX AGRO-BIO TECH – UNIVERSITY OF LIÈGE, 223P., 29 TABL., 7 FIG.

ABSTRACT

Animal breeding programs are designed to genetically improve livestock populations over many generations to enhance farm sustainability and competitiveness. Genetic improvement is achieved by selecting genetically superior animals, based on estimated breeding values (EBV), to be the parents of the next generation. These EBV are calculated by solving mixed model equations characterizing appropriate statistical genetic evaluation models. To guarantee effective genetic selection, genetic evaluation models must be tailored to the specific characteristics of the traits and population under evaluation. This PhD thesis focused on the development of genetic evaluation models suitable for categorical maternally influenced traits and for multi-breed populations. Appropriate genetic animal models were developed and assessed: (1) for two categorical maternally influenced traits based on calving ease scores from Walloon Holstein dairy cattle and on lamb survival data from a New Zealand sheep population; (2) for two multi-breed populations based on milk yield records of New Zealand purebred and crossbred dairy cattle, and on purebred and crossbred calving ease scores from Walloon Belgian Blue and Holstein cattle. Results showed that (1) fitting maternal effects was required to avoid biasing the EBV, and there was no clear advantage in using non-linear mixed models instead of linear mixed models for the genetic analysis of the two categorical maternal traits studied; (2) breed-dependent EBV could be estimated using the proposed multi-breed models, and that combining purebred and crossbred data had a positive influence on the accuracy of the EBV of purebred animals. Finally, part of the research presented in this thesis contributed to the development of the genetic evaluation systems currently used in Walloon Region of Belgium and in New Zealand.

VANDERICK SYLVIE. (2017). AMÉLIORATION DES SYSTÈMES D'ÉVALUATION GÉNÉTIQUE POUR DES CARACTÈRES MATERNELS ET DES POPULATIONS MULTI-RACES D'ANIMAUX D'ÉLEVAGE. (THÈSE DE DOCTORAT EN ANGLAIS). GEMBLOUX, BELGIQUE, GEMBLOUX AGRO-BIO TECH – UNIVERSITÉ DE LIÈGE, 223P., 29 TABL., 7 FIG.

RÉSUMÉ

Les programmes de sélection animale sont conçus pour améliorer génétiquement les animaux d'élevage au fil des générations afin de renforcer la durabilité et la compétitivité des exploitations agricoles. L'amélioration génétique est obtenue en sélectionnant – sur base de valeurs d'élevage estimées – des animaux génétiquement supérieurs comme parents de la prochaine génération. Ces valeurs d'élevage estimées sont calculées en résolvant les équations du modèle mixte caractérisant des modèles statistiques appropriés d'évaluation génétique. Afin de garantir une sélection génétique efficace, les modèles d'évaluation génétique doivent être adaptés aux caractéristiques spécifiques des caractères et de la population à évaluer. Cette thèse de doctorat s'est concentrée sur le développement de modèles d'évaluation génétique adaptés aux caractères maternels et aux populations multi-races d'animaux d'élevage. Des modèles d'évaluation génétique ont été développés et testés: (1) pour deux caractères discrets maternellement influencés à partir de données relatives à la facilité de vêlage des bovins laitiers wallons de race Holstein, et de données relatives à la survie des agneaux d'une population de moutons de Nouvelle-Zélande; (2) pour deux populations multi-races à partir de données de production laitière de bovins laitiers purs et croisés de Nouvelle-Zélande, et de données relatives à la facilité de vêlage de bovins wallons de race Holstein et Blanc Bleu Belge et de leurs croisés. Les résultats ont montré: (1) qu'il était nécessaire de tenir compte des effets maternels afin de ne pas biaiser les valeurs d'élevage estimées, et qu'il n'y avait pas d'avantage particulier à utiliser des modèles mixtes non-linéaires plutôt que des modèles mixtes linéaires pour l'évaluation génétique des deux caractères discrets maternels étudiés; (2) que des valeurs d'élevage dépendantes de la race pouvaient être estimées au moyen des modèles multi-races développés, et que la combinaison des données d'animaux purs et croisés avait une influence positive sur la précision des valeurs d'élevage des animaux purs. Enfin, il est important de signaler qu'une partie de la recherche présentée dans cette thèse a contribué au développement de systèmes d'évaluation génétique utilisés en Région Wallonne et en Nouvelle-Zélande.

ACKNOWLEDGEMENTS

After several years and an intensive period of several months, today is the day: writing these acknowledgements is the finishing touch on my PhD thesis. I would like to take this opportunity to thank the people that directly or indirectly have helped and supported me along this challenging period.

Firstly, I would like to express my appreciation and thanks to my promoter, Prof. Nicolas Gengler for his support, help and guidance through this work, for his motivation and immense knowledge. Thank you for offering me the opportunities to travel abroad and meet many people from the quantitative genetics community.

Besides my promoter, I would like to extend my acknowledgements to the rest of my thesis committee: Benoit Auvray (University of Otago, New Zealand), Carlo Bertozzi (Walloon Breeding Association, Belgium) and Prof. Yves Beckers from Gembloux Agro-Bio Tech, University of Liège, for their helpful comments and suggestions, as well as for their encouragement. Moreover, I would like to thank all the members of the jury: Prof. Marie-Laure Fauconnier (President of the jury), Dr. Hélène Soyeurt and Dr. Ludivine Lassois from Gembloux Agro-Bio Tech, University of Liège.

I wish to thank the awé asbl (Walloon Breeding Association, Ciney, Belgium) for the technical support, especially Alain Gillon and Géry Glorieux who have always been available to answer my questions about the data and pedigree files.

I also thank the co-authors of the papers included in my thesis for their valuable input, advice and thoughtful contributions: Benoit Auvray, Dr. Ken Dodds, Dr. Julie Everett-Hincks, Dr. Pierre Faux, Alain Gillon, Géry Glorieux, Dr. Bevin Harris, Patrick Mayeres, Dr. Rodrigo Reis Mota, Dr. Sheryl Anne Newman, Dr. Jenny Pryce, and Thibault Troch.

Julie, Benoit, I am so grateful that you took me in under your wing during 6 months to work with the lamb survival team. I have learned a lot by working with you, and I have wonderful memories from the time spent there.

Special thanks go to a bunch of people, Aurélie Lainé, Benoit Auvray, Hedi Hammami and Rodrigo Reis Mota, for the time they spent to read and correct the different drafts of the chapters of my PhD thesis. I am also grateful to Frédéric Colinet for his valuable administrative advices as well as for his organizational tips.

And finally, my acknowledgements go to the Public Service of Wallonia (Service Public de Wallonie, Direction générale opérationnelle de l’Agriculture, des Ressources naturelles et de l’Environnement (SPW-DGO3)) for its financial support through the “*Convention de collaboration technique et scientifique dans le cadre des évaluations génétiques bovines laitières pour la Wallonie*” ; and to the University of Liège for the grant provided for my scientific stay of 6 months at AgResearch Limited, Invermay Agricultural Centre, New Zealand.

REMERCIEMENTS

Après les remerciements officiels, voici ceux plus officieux…

Benoit et Kathleen, merci pour votre hospitalité et votre convivialité durant notre séjour au pays du long nuage blanc. Merci pour votre aide et vos conseils dans différents domaines : logistique, administratif, touristique, gastronomique… Grâce à vous, nous avons pu profiter pleinement de nos 6 mois (bien trop courts) dans l'île du Sud. Merci à Adrien, Elise et Olive de nous avoir acceptés dans leurs quotidiens et surtout d'avoir intégré Ysaline à leurs jeux. Merci à Elise d'avoir aidé Ysaline dans ses débuts à la Kindy.

Un grand merci également aux collègues, stagiaires et étudiants qui ont fait et font vivre le service de « Zootechnie ». Geneviève et Cécile, merci beaucoup pour votre disponibilité, votre soutien et votre amitié. Yves, merci pour tes encouragements et tes conseils tout au long de ces années de thèse, particulièrement durant ces derniers mois. Fred, merci d'avoir été là, d'être là… je t'avais dit que je t'écrirai une « tartine » dans ces remerciements et puis au final je t'ai tout dit lors de ton départ… bref tu sais ! Hedi, « *chokran jazilan 3aychek* » pour ton aide, tes conseils et tes idées, ainsi que pour nos (longues) discussions sur les enfants, le travail, la vie. Rodrigo, « *muito obrigada* » for taking the time to read, correct and improve several parts of this thesis. Pierre, merci de m'avoir convaincue de terminer cette thèse … voyons maintenant ce que l'avenir (professionnel) me réserve ! Aurélie, Alain, Catherine, Marie, Jérémie, Marie-Laure, Valérie, Hana, Laura, Elisabeth, Hélène, Amaury … merci à vous pour tous ces moments que nous avons vécu ensemble au sein de l'équipe de Nicolas, que ça soit à Gembloux ou à Uppsala, Belo Horizonte, Kuopio, Dublin, Paris, Barcelone, Riga, Leipzig, Wageningen, Athens, Berlin … Je n'oublie bien sûr pas mes deux tfistes ! Johan et Thibault, merci pour votre aide dans le développement de l'évaluation génétique wallonne de la facilité de vêlage.

Elisabeth, Ellen et Coraline ainsi que vos moitiés (Grégoire, Benoit, Denis)… merci de votre amitié, de votre soutien et de vos encouragements.

Pour conclure, je veux souligner que l'aboutissement de cette thèse n'aurait pas été possible sans l'amour, l'appui et les encouragements de mon mari, de mes deux merveilleuses princesses et de ma maman. Vous avez dû supporter pendant plusieurs mois mes angoisses, mes doutes, mes larmes, ma mauvaise humeur, ma fatigue ainsi que mes absences durant plusieurs soirées, weekends et vacances. Merci pour tout, je vous aime. Merci également à mon frère et ma belle-sœur d'être toujours là pour moi… « ça y est je l'ai enfin finie cette thèse !!! ». Petite dédicace spéciale à mon *être diabolique.*

Au moment de terminer d'écrire ces quelques lignes, mon cœur ne peut s'empêcher de se serrer à l'idée que mon papa aurait souhaité être présent pour vivre ce moment avec moi, avec nous. Merci papa d'avoir toujours été là également pour moi… tu aurais été fier de moi, non tu es fier de moi !

TABLE OF CONTENTS

CHAPTER I. GENERAL INTRODUCTION

CONTEXT

Animal breeding aims to genetically improve livestock populations by selecting and mating the best individuals as parents of the next generations in order to increase sustainability and long-term economic profitability of livestock production (*e.g.* reducing the production costs, improving animal health and welfare). A large variety of statistical techniques, strategies and methods have been developed to achieve this goal.

Genetic selection based on estimated breeding values (EBV) is the most important contribution to performance improvement of livestock animals (Havenstein *et al.*, 2003). These EBV, which are traditionally based on the knowledge of phenotypic observations and pedigree information, make possible the comparison and the ranking of animals between them, thereby to select the ones with the highest genetic merit.

There has been considerable work in the last decades to adapt genetic evaluation models, computational methods and software to the specific characteristics of each trait/population evaluated. Also, since the early 2000s, the inclusion of molecular information, such as single nucleotide polymorphism (SNP), has become increasingly important in animal breeding. Currently, most developments are focusing on an approach called “genomic selection” or more properly “genomic prediction” (Meuwissen *et al.*, 2001; Hayes *et al.*, 2009; VanRaden *et al.*, 2011). In this method dense SNP panels allow to link also phenotypes and SNP genotypes to predict the genetic merit of animals, thus leading to a new approach to predict genetic merits based also on molecular information combined with phenotypes and pedigrees.

However, these advances do not change the fundamental issue that genetic evaluation models and methods need to be adjusted to existing needs of breeders and emerging issues. Some examples where practical solutions are required are the genetic evaluation of discrete traits (*e.g.* disease susceptibility, litter size, reproductive traits), the genetic evaluation of maternally influenced traits (*e.g.* birth or weaning weight, calving ease, viability), and the genetic evaluation of animals in multi-breed and crossbred populations. The present document focuses particularly on the development of genetic evaluation systems for genetic evaluation of maternally influenced traits, as well as for genetic evaluation of animals in multi-breed populations.

Maternally influenced traits are traits that are influenced not only by the individual itself, but also by its mother. The effects linked to the recorded individual are, in general, called direct effects. Direct effects can be due to the genes, but can also be non-genetic (*e.g.* animal specific environment). The effects linked to the mother are, in general, called maternal effects and can be considered as the genetic and non-genetic ability of mother to provide, in a broad sense, a suitable environment to her offspring (Willham, 1972). Non-genetic effects are the not transmitted part of the contribution (environment) by the mother. These effects are often called maternal environmental effects, even if they only represent the non-genetic part of the maternal environment. However, the part that can be transmitted is called maternal genetic effects. Many traits of interest in livestock production, such as for example calving ease in cattle, lamb survival to weaning, or early growth traits in beef and swine, are affected by maternal effects (Dematawena and Berger, 1997; Hawkins and Wu, 2011). Consequently, genetic evaluation systems of maternally influenced traits have to account for these maternal effects, in addition to those associated to the individual itself (*i.e.* direct effects) in order to estimate unbiased EBV needed by breeders and producers for their selection and mating decisions.

Usually, genetic improvement programs of livestock animals may involve two different breeding strategies that might be combined: 1) selection of best animals within breeds or populations, *i.e.* pure (or straight) breeding systems, and 2) using the best breeds or breed combination through crossbreeding systems. Very different crossbreeding systems exist, some focusing on creating animals with improved germplasms to be transmitted to the next generations, others to produce production animals or combinations of both. Crossbreeding is widely used in pig, poultry and beef cattle production systems (Yang and Jiang, 2005; Dufrasne, 2014; Theunissen *et al.*, 2014). These production systems rely heavily on heterosis and breed complementarity effects to improve productivity and efficiency of production in the crossbred offspring (Hansen, 2006). Crossbreeding of *Bos taurus* dairy breeds with local *Bos indicus* cattle is also a well-documented strategy to enhance milk production in tropical climates (Galukande *et al.*, 2013). In contrast, in temperate climates, crossbreeding in dairy cattle has not been widespread, except in New Zealand where about 46% of cows are crossbreds (New Zealand Dairy Statistics 2014-15). The rationale behind the fact that pure breeding systems are most commonly used for dairy cattle in most temperate countries is due in large part because of the increasing global dominance of Holstein cattle in the dairy industry (Weigel and Barlass, 2003). However, for several years, trends in milk pricing, as well as increasing health, fertility, calving and longevity performance concerns for most of the high-producing dairy breeds (*e.g.* Heins *et al.*, 2006; Pritchard *et al.*, 2013), have led dairy breeders and producers to become more interested in the use of crossbreeding (Sørensen *et al.*, 2008). Furthermore, increasing levels of inbreeding within those breeds (Croquet *et al.*, 2006; Miglior *et al.*, 2008) have also contributed to the growing interest in crossbreeding across the world (Sørensen *et al.*, 2008). Most genetic evaluation systems in dairy cattle compare animals only within breed (Interbull, 2016), accordingly based only on purebred information and ignoring information from crossbred animals. However, traits that are evaluated in purebred populations may be genetically different from traits at the crossbred level because the genetic correlations between purebred and crossbred performance are usually estimated to be less than one (*e.g.* Zumbach *et al.*, 2006; Mulder *et al.*, 2016). Thus, the genetic merit estimated only on purebred performance is not a good predictor of performance in crossbreeding. Therefore, genetic evaluation systems need to be adapted to combine purebred and crossbred data to improve selection of purebred parents for performance in crossbreeding (Wei and van der Werf, 1994; VanRaden *et al.*, 2007).

THIS THESIS

The overall aim of this thesis is to present genetic evaluation systems developed to improve genetic selection of livestock animals:

1. For categorical maternally influenced traits;
2. In multi-breed populations composed of purebred and crossbred animals.

This thesis is a compilation of published scientific papers and is structured as follows.

First, an overview of the mixed model methodology used in the context of animal breeding and genetic selection is presented (Chapter II).

Then, Chapter III focuses on the development of genetic evaluation systems for categorical maternally influenced traits using appropriate maternal models based on 1) calving ease scores from Walloon Holstein dairy cattle (Chapters III.1 and III.2) and 2) data of lamb survival to weaning from New Zealand sheep (Chapters III.3).

In Chapter III.1, linear and threshold maternal animal models to analyse calving ease scores from the Holstein dairy cattle in the Walloon Region of Belgium are firstly assessed and compared in terms of predictive ability while examining and discussing the correlation between direct and maternal additive genetic effects. Then, Chapter III.2 focuses on the implementation in routine of the genetic evaluation to predict the EBV (*i.e.* genetic merit) of Walloon Holstein dairy cattle for calving ease performance.

In Chapter III.3, different maternal animal models to analyse lamb survival to weaning data are compared in terms of predictive ability. Thus, the benefits of using generalized linear models rather than linear models as well as including maternal environmental effects in the genetic evaluation model are explored.

Then, Chapter IV focuses on the development of genetic evaluation models adapted to multi-breed populations composed of purebred and crossbred animals using 1) first-lactation milk yield data of New Zealand dairy cattle where crossbreeding is common (Chapter IV.1) and 2) calving ease scores from purebred and crossbred animals of both dominant cattle breeds in the Walloon Region of Belgium: the Holstein and Belgian Blue breeds (Chapter IV.2).

In Chapter IV.1, a multi-breed model based on a random regression test-day model combining purebred and crossbred data is proposed to model different additive genetic effects according to breed composition. Genetic parameters estimated by this multi-breed model are compared to estimates from within breeds using a single-breed random regression test-day model. Furthermore, usefulness of this kind of complicated random regression test-day model is investigated.

Chapter IV.2 examines the usefulness and feasibility of a joint genetic evaluation system for calving ease combining data from purebred and crossbred animals. For this purpose, two multi-breed random regression maternal models using different functions of breed proportions as random regression predictors are assessed and compared in terms of goodness of fit. Likewise, a comparison of genetic parameters estimated within and across breeds is carried out. Furthermore, the potential gain of reliability for EBV for purebred sires when combining purebred and crossbred data is checked.

Finally, Chapter V presents a general discussion of the results obtained through previous chapters. Several aspects related to the accuracy and efficiency of genetic selection based on EBV are discussed: 1) the quantity and quality of phenotypic data that are the basis of any genetic/genomic evaluation systems, 2) the adequacy of the statistical genetic evaluation model to guarantee an accurate estimation of genetic parameters and EBV. Thus, models for categorical traits, maternally influenced traits and for multi-breed populations are discussed. Also, general conclusions and future perspective of research are addressed.

FRAMEWORK

This thesis was initiated in the context of the development and implementation of the genetic evaluation systems of Walloon dairy cattle under the framework of the “Convention de collaboration technique et scientifique dans le cadre des évaluations génétiques bovines laitières pour la Wallonie”, jointly managed by the Numerical Genetics, Genomics and Modeling Group from Agriculture, Bio-engineering and Chemistry Department of Gembloux Agro-Bio Tech, University of Liège (ULg-GxABT; Gembloux, Belgium) and the Walloon Breeding Association (awé asbl, Ciney, Belgium) and fully funded by the Public Service of Wallonia (Service Public de Wallonie – Direction Générale Opérationnelle de l’Agriculture, des Ressources naturelles et de l’Environnement). The aim of this convention was to develop management and selection tools, chiefly through the estimation of the genetic merit of animals, to support dairy farmers in their breeding decisions. The main part of the research was undertaken at Gembloux Agro-Bio Tech, under the supervision of Prof. Nicolas Gengler, on data collected and provided by the awé. A part of the research was realised at the AgResearch Limited, Invermay Agricultural Centre (Mosgiel, New Zealand), under the supervision of Dr. Julie Everett-Hincks and Benoit Auvray, on data provided by Sheep Improvement Limited (Christchurch, New Zealand). This research was partly funded by a grant from the ULg. Lastly, our collaboration with the Livestock Improvement Corporation (LIC, Hamilton, New Zealand) allowed the work on New Zealand dairy cattle data.

LITERATURE CITED

Croquet, C., P. Mayeres, A. Gillon, S. Vanderick, and N. Gengler. 2006. Inbreeding depression for global and partial economic indexes, production, type, and functional traits. J. Dairy Sci. 89:2257–2267.

Dematawena, C. M. B., and P. J. Berger. 1997. Effect of dystocia on yield, fertility, and cow losses and an economic evaluation of dystocia scores for Holsteins. J. Dairy Sci. 80:754–761.

Dufrasne, M. 2014. Genetic improvement of pig sire lines for production performances in crossbreeding [Ph.D. Thesis]. Université de Liège, GxABT, Gembloux, Belgique.

Galukande, E., H. Mulindwa, M. Wurzinger, R. Roschinsky, A. O. Mwai, and J. Sölkner. 2013. Cross-breeding cattle for milk production in the tropics: achievements, challenges and opportunities. Anim. Genet. Resour. Génétiques Anim. Genéticos Anim. 52:111–125.

Hansen, L. 2006. Monitoring the worldwide genetic supply for dairy cattle with emphasis on managing crossbreeding and inbreeding. In: Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006. Instituto Prociência. p. 1–1.

Havenstein, G., P. Ferket, and M. Qureshi. 2003. Carcass composition and yield of 1957 versus 2001 broilers when fed representative 1957 and 2001 broiler diets. Poult. Sci. 82:1509–1518.

Hawkins, B. W., and S. Wu. 2011. Understanding the variability in sheep and beef farm profitability. Proc. N. Z. Grassl. Assoc. 73:23–26.

Hayes, B. J., P. J. Bowman, A. J. Chamberlain, and M. E. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. J. Dairy Sci. 92:433–443.

Heins, B., L. Hansen, and A. Seykora. 2006. Calving difficulty and stillbirths of pure Holsteins versus crossbreds of Holstein with Normande, Montbeliarde, and Scandinavian Red. J. Dairy Sci. 89:2805–2810.

Interbull. 2016. Description of National Genetic Evaluations Systems for dairy cattle traits as applied in different Interbull member countries. Int. Bull Eval. Serv. Off. Website. Available from: http://www.interbull.org/ib/geforms

Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. Genetics 157:1819–1829.

Miglior, F., B. Van Doormaal, and G. Kistemaker. 2008. Phenotypic analysis of inbreeding depression for traits measured in Canadian dairy cattle breeds. In: 2008 Annual Meeting of the Canadian Society of Animal Science (CSAS). Canada.

Mulder, H. A., J. Visscher, and J. Fablet. 2016. Estimating the purebred–crossbred genetic correlation for uniformity of eggshell color in laying hens. Genet. Sel. Evol. 48. Available from: http://gsejournal.biomedcentral.com/articles/10.1186/s12711-016-0212-2

New Zealand Dairy Statistics 2014-15. prepared by LIC and Dairy NZ, Hamilton, New Zealand. Available from: www.dairynz.co.nz/dairystatistics

Pritchard, T., M. Coffey, R. Mrode, and E. Wall. 2013. Genetic parameters for production, health, fertility and longevity traits in dairy cows. Animal 7:34–46.

Sørensen, M. K., E. Norberg, J. Pedersen, and L. G. Christensen. 2008. Invited Review: Crossbreeding in Dairy Cattle: A Danish Perspective. J. Dairy Sci. 91:4116–4128.

Theunissen, A., M. Scholtz, F. Neser, and M. MacNeil. 2014. Crossbreeding to increase beef production: Additive and non-additive effects on fitness traits. South Afr. J. Anim. Sci. 44:335–349.

VanRaden, P. M., J. R. O’Connell, G. R. Wiggans, and K. A. Weigel. 2011. Genomic evaluations with many more genotypes. Genet. Sel. Evol. 43:1.

VanRaden, P. M., M. E. Tooker, J. B. Cole, G. R. Wiggans, and J. H. Megonigal. 2007. Genetic Evaluations for Mixed-Breed Populations. J. Dairy Sci. 90:2434–2441.

Wei, M., and J. H. J. van der Werf. 1994. Maximizing genetic response in crossbreds using both purebred and crossbred information. Anim. Prod. 59:401–413.

Weigel, K. A., and K. A. Barlass. 2003. Results of a producer survey regarding crossbreeding on US dairy farms. J. Dairy Sci. 86:4148–4154.

Willham, R. L. 1972. The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. J. Anim. Sci. 35:1288–1293.

Yang, N., and R.-S. Jiang. 2005. Recent advances in breeding for quality chickens. Worlds Poult. Sci. J. 61:373–381.

Zumbach, B., I. Misztal, S. Tsuruta, J. Holl, W. Herring, and T. Long. 2006. Genetic correlations between two strains of Durocs and crossbreds from differing production environments for slaughter traits. J. Anim. Sci. 85:901–908.

CHAPTER II. OVERVIEW OF MIXED MODEL METHODOLOGY USED IN ANIMAL BREEDING AND GENETICS

CHAPTER II. OVERVIEW OF MIXED MODEL METHODOLOGY USED IN ANIMAL BREEDING AND GENETICS

OUTLINE

Animal breeding is often concerned with variation in traits of interest of offspring as supposed inherited from their sires and dams. Researchers in the field of quantitative genetics have been mainly involved with on-farm animal recording procedures gathering data that have hierarchical structure associated with various levels of variation. One of the most important contributions to performance improvement is the genetic selection based on estimated breeding values, which are the solutions of mixed model equations. Obviously, the mixed model structure depends on the trait to be evaluated and the population to be analysed; thus enabling to provide accurate breeding values for a given trait in a given population. Therefore, the purpose of this chapter is to review general concepts of mixed models as applied to animal breeding and genetic selection.

INTRODUCTION

Animal breeding is a powerful tool that allows breeders and also producers to select superior animals to reduce production costs (*e.g.* milk, meat, eggs, wool), to improve animal health and welfare, thus to increase farm sustainability over the generations by improving the genetic merit of livestock. Genetic selection makes a most important contribution to performance improvement (Havenstein *et al.*, 2003). Animals with the highest genetic merit are selected on the basis of accurate values, called "estimated breeding values" (EBV) or "genetic values", which allow the comparison and the classification of animals. Today, these EBV are mostly derived from the solutions of mixed model equations (MME) established by Henderson in 1973.

THE GENERAL LINEAR MIXED MODEL

Henderson adapted theories of linear mixed models to quantitative genetic. The term "mixed model" refers to the use of both fixed and random effects in the same analysis model. In many genetic applications, a linear mixed model is assumed and can be represented in matrix notation as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \qquad (1)$$

in which $\mathbf{y}$ is the vector of phenotypic observations on the traits and animals of interest; $\mathbf{b}$ is the vector of unknown fixed effects that influence the phenotypic observations; $\mathbf{u}$ is the vector of unknown random effects that influence the phenotypic observations; $\mathbf{e}$ is the vector of the unknown residual effects; and $\mathbf{X}$ and $\mathbf{Z}$ are the known incidence matrices relating observations to corresponding fixed and random effects, respectively.

The fixed and random effects allow modelling the expectations and the variability of the phenotypic observations ($\mathbf{y}$) respectively. Thus, the expectations and variances of $\mathbf{y}$ are assumed to be:

$$E(\mathbf{y}) = E(\mathbf{Xb}) \qquad \text{as } E(\mathbf{u}) = E(\mathbf{e}) = 0 \tag{2}$$

$$\begin{aligned}
\operatorname{var}(\mathbf{y}) &= \operatorname{var}(\mathbf{Xb} + \mathbf{Zu} + \mathbf{e}) & & \\
&= \operatorname{var}(\mathbf{Xb}) + \operatorname{var}(\mathbf{Zu} + \mathbf{e}) & & \text{as } \operatorname{var}(\mathbf{Xb}) = 0 \\
&= \mathbf{Z}\operatorname{var}(\mathbf{u})\mathbf{Z}' + \operatorname{var}(\mathbf{e}) + \mathbf{Z}\operatorname{cov}(\mathbf{u},\mathbf{e}) + \operatorname{cov}(\mathbf{e},\mathbf{u})\mathbf{Z}' & & \\
&= \mathbf{ZGZ}' + \mathbf{R} + \mathbf{Z}\operatorname{cov}(\mathbf{u},\mathbf{e}) + \operatorname{cov}(\mathbf{e},\mathbf{u})\mathbf{Z}' & & \text{as } \operatorname{var}(\mathbf{u}) = \mathbf{G} \text{ and } \operatorname{var}(\mathbf{e}) = \mathbf{R} \\
&= \mathbf{ZGZ}' + \mathbf{R} = \mathbf{V} & & \text{as } \operatorname{cov}(\mathbf{u},\mathbf{e}) = \operatorname{cov}(\mathbf{e},\mathbf{u}) = 0
\end{aligned} \tag{3}$$

The resolution of model (1) using Best Linear Unbiased Estimator (**BLUE**) and Best Linear Unbiased Prediction (**BLUP**) equations provides the following solutions for the fixed ($\hat{\mathbf{b}}$) and random effects $(\hat{\mathbf{u}})$ (Henderson, 1973; Henderson, 1975):

$$\hat{\mathbf{b}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \tag{4}$$

$$\hat{\mathbf{u}} = \mathbf{GZ}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}) \tag{5}$$

The solutions of equations (4) and (5) require the inverse of the covariance matrix $\mathbf{V}$, which is usually too large to be inverted. However, Henderson (1975) offered a method for jointly obtaining $\hat{\mathbf{b}}$ and $\hat{\mathbf{u}}$ in the form of his MME. Thus; these equations (6) allow estimating solutions of fixed effects and predicting solutions for random effects simultaneously without the need for computing $\mathbf{V}^{-1}$. The MME for model (1) are:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \tag{6}$$

where $\mathbf{R}$ is the covariance matrix for residual effects and $\mathbf{G}$ is the covariance matrix for random effects as shown previously in (3). These equations are of order equal to the number of elements in $\mathbf{b}$ and $\mathbf{u}$, which is usually much less than the number of elements in $\mathbf{y}$, and therefore, are more practical to solve. Also, these equations require the inverse of $\mathbf{R}$ rather than $\mathbf{V}$, both of which are of the same order, but $\mathbf{R}$ is generally diagonal or has a more simple structure than $\mathbf{V}$. Likewise, the inverse of $\mathbf{G}$ is needed, which is of order equal to the number of elements in $\mathbf{u}$. The ability to compute the inverse of $\mathbf{G}$ depends on the model and the definition of the random effects.

The MME structure displayed in (6) is expandable in many ways. Both vectors $\mathbf{b}$ and $\mathbf{u}$ may contain one or more factors depending on the situation. The vector $\mathbf{b}$ may include several fixed effects such as age at birth, herd-year-season of birth, number of times milked, size of litter, breed composition and possibly other identifiable environmental factors. Likewise, vector $\mathbf{u}$ may include several random effects, such as animal and permanent environmental effects (*e.g.* a repeatability model) or such as animal and maternal effects (*e.g.* a maternal model). Of course, this leads to a more complex structure of $\mathbf{G}$, so more difficulties to invert $\mathbf{G}$. In the same way, the vector $\mathbf{y}$ could contain phenotypic observations for several (correlated) traits, and consequently $\mathbf{u}$ would contain several animal and environmental effects referring to the different traits as is the case in a multiple-trait model. In addition, the $\mathbf{R}$ matrix could reflect correlations between residuals or different residual variances for different groups of observations (*e.g.* multiple-trait model). Therefore, in the establishment of a linear mixed model, not only fixed effects have to be defined, but also the covariance structure of the random and residual effects.

SIMPLE ANIMAL MODEL

The simple linear mixed model used in animal breeding is the single-trait animal model with only one random effect; hereafter this simple model is called “the simple animal model”. The term “animal model” can be used when there is an animal effect as random effect involved in phenotypic observations, meaning that an EBV (or a genetic value) is fitted for each animal. The term “single-trait” refers to the fact that animals have only phenotypic observations on one trait. Thus, in the simple animal model, only fixed and animal random effects influence the phenotypic observations from a single trait. The principle of this model is to apply MME to include all relatives – with or without phenotypic observations – to evaluate simultaneously dams and sires enhancing the accuracy of genetic merit of animals. This is possible by the use of the additive genetic (or numerator) relationship matrix. The simple animal model can be expressed in matrix notation as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e} \qquad (7)$$

In this model, assuming purely additive gene action (*i.e.* assuming that only the genetic parts are passed from parents to progeny), the performance of an animal is described genetically according to the value of its additive genetic effect, *i.e.* its EBV. The unknown vector $\mathbf{a}$ in (7) contains the animals’ additive effects, which are fitted as random effects. The vectors $\mathbf{y}$, $\mathbf{b}$, and $\mathbf{e}$ as well as the incidence matrices $\mathbf{X}$ and $\mathbf{Z}$ were previously described in (1). The variances of the random and residual effects from model (7) are assumed to be:

$$\text{var}\begin{bmatrix}\mathbf{a}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{A}\sigma_a^2 & 0\\ 0 & \mathbf{I}\sigma_e^2\end{bmatrix} = \begin{bmatrix}\mathbf{G} & 0\\ 0 & \mathbf{R}\end{bmatrix} \qquad (8)$$

where $\mathbf{G}$ is the covariance matrix for the additive effects and is equal to $\mathbf{A}\sigma_a^2$ being $\mathbf{A}$ the additive relationship matrix and σ_a^2 the variance due to the additive effects; $\mathbf{R}$ is the residual covariance matrix, in which residual effects are, in general, assumed independently distributed with the variance σ_e^2; and $\mathbf{I}$ is an identity matrix. Moreover, additive effects and residual effects are usually assumed uncorrelated.

By factoring out $\mathbf{R}^{-1}$ from both sides of the MME (6) and replacing $\hat{\mathbf{u}}$ by $\hat{\mathbf{a}}$, the MME for the simple animal model (7) are:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{Z} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}\lambda_a \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \end{bmatrix} \qquad (9)$$

with $\lambda_a = \dfrac{\sigma_e^2}{\sigma_a^2}$.

The vector $\hat{\mathbf{a}}$ in (9) contains solutions for the additive effects, *i.e.* the EBV on which breeders can operate for selecting animals with the best genetic merit. The equations in (9) require the inverse of $\mathbf{A}$. An efficient and feasible method, even for large populations, to compute directly the inverse of $\mathbf{A}$ (*i.e.* $\mathbf{A}^{-1}$) have been presented by (Henderson, 1976) and (Quaas, 1976).

Of course, this simple animal model is a simplified approach and can be easily extended to include animals' additive random effects for several traits (*e.g.* multiple-trait animal model), maternal random effects (*e.g.* maternal animal model), non-additive genetic random effects, such as dominance and epistasis (*e.g.* non-additive animal model), repeated observations on individuals (*e.g.* repeatability animal model) and so on.

ANIMAL MODEL WITH GENETIC GROUPS

As stated previously, an animal model requires that complete pedigrees are known for all animals with or without observations. However, pedigree information on each animal may not be traceable back to a base population due to lack of recording and/or movement of animals from one owner to another (in the same or different countries). Thus, in any pedigree file there are always animals whose parents are unknown or missing.

In an animal model, the use of relationship matrix **A** considers these animals with unknown parents as being sampled from arbitrarily defined base population. Members of this base population are assumed to originate from a single large population with an average EBV of zero and a common additive variance of σ_a^2. In this population, all the individuals are randomly mating, such that they are unrelated to each other and non-inbred. However, the assumption that all animals with unknown parents come from one single large population is rarely true in practical applications. Indeed, if we know that animals without known parents were actually sampled from populations with different genetic means because of, for instance, selection intensity or parents' country of origin, then the model must account for the subpopulation structure and allow for different expected values of base population animals. A way to deal with this is to create genetic groups and include their effects in the model. Doing this avoids the assumption that animals with unknown parents all come from a single large base population and allows taking into account differences in genetic means between populations.

Several grouping strategies were proposed (*e.g.* Thompson, 1979; Famula *et al.*, 1983) but the grouping strategy developed by Robinson (1986), Quaas (1988) and Westell *et al.* (1988) has generally been adopted. In this strategy, a genetic group for each animal is derived from the genetic group effects of the animal's ancestors. For each known animal with unknown parents, phantom parents without observations are created and are assigned to appropriate genetic groups. Obviously, there are many ways to assign phantom parents to genetic groups. They can be assigned for instance by sex, birth year (time), country of origin, breed, and intensity of selection (selection path). The proper definition of the genetic groups as well as their fitting as fixed or random effects will depend on the knowledge of the data and the intended purpose of the analysis.

If genetic group effects are added to the simple animal model (7), model becomes as:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{ZQg} + \mathbf{Za} + \mathbf{e} \qquad (10)$$

where **g** is the vector of genetic group effects, which represent the average genetic merit of the phantom animals selected to be parents to their descendants that do have phenotypic observations available, and $\mathbf{Q}$ is a matrix containing the proportion of gene contribution from genetic groups to the animals. The vectors $\mathbf{y}$, $\mathbf{b}$, $\mathbf{a}$ and $\mathbf{e}$ and matrices $\mathbf{X}$ and $\mathbf{Z}$ were described in (7).

With model (10) the EBV of an animal is equal to the mean of its group plus its individually specific deviation from the group, *i.e.* $\text{vector of EBVs} = \hat{\mathbf{u}} = \mathbf{Q}\hat{\mathbf{g}} + \hat{\mathbf{a}}$. From model (10), the corresponding MME are:

$$\begin{bmatrix} \mathbf{X'X} & \mathbf{X'ZQ} & \mathbf{X'Z} \\ \mathbf{Q'Z'X} & \mathbf{Q'Z'ZQ} & \mathbf{Q'Z'Z} \\ \mathbf{Z'X} & \mathbf{Z'ZQ} & \mathbf{Z'Z} + \mathbf{A}^{-1}\lambda_a \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{g}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'y} \\ \mathbf{Q'Z'y} \\ \mathbf{Z'y} \end{bmatrix} \quad (11)$$

with $\lambda_a = \dfrac{\sigma_e^2}{\sigma_a^2}$.

Thanks to QP transformation (Quaas and Pollak, 1981), the MME in (11) can be modified so that $\mathbf{Q}\hat{\mathbf{g}} + \hat{\mathbf{a}}$ can be computed directly as the following MME:

$$\begin{bmatrix} \mathbf{X'X} & \mathbf{0} & \mathbf{X'Z} \\ \mathbf{0} & \mathbf{Q'A}^{-1}\mathbf{Q}\lambda_a & -\mathbf{Q'A}^{-1}\lambda_a \\ \mathbf{Z'X} & -\mathbf{A}^{-1}\mathbf{Q}\lambda_a & \mathbf{Z'Z} + \mathbf{A}^{-1}\lambda_a \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{g}} \\ \mathbf{Q}\hat{\mathbf{g}} + \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'y} \\ \mathbf{0} \\ \mathbf{Z'y} \end{bmatrix} \quad (12)$$

Furthermore, procedures developed by Westell *et al.* (1988) enable to write directly the elements of the inverse of relationship matrix including genetic groups (*i.e.* $\mathbf{Q'A}^{-1}\mathbf{Q}$, and $-\mathbf{Q'A}^{-1}$) to be computed directly using Henderson's rules (Henderson, 1976).

REPEATABILITY ANIMAL MODEL

When multiple phenotypic observations on the same trait are recorded on the same animal, such as litter size in successive pregnancies or milk yield records in successive lactations, a second random effect needs to be introduced, for each animal with observations, in the simple animal model (7) to account for shared environmental effects. Hence, the model becomes,

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Z}_{pe}\mathbf{p} + \mathbf{e} \qquad (13)$$

where $\mathbf{p}$ is the vector of permanent environmental effects and non-additive genetic effects; and $\mathbf{Z}_{pe}$ is the known incidence matrix relating phenotypic observations to permanent environmental effects. It is important to emphasise that $\mathbf{Z}_{pe}$ = $\mathbf{Z}$ if each animal has a permanent environmental effect as well as an additive effect. If there are animals without observations and if their EBV are also desired then $\mathbf{Z}_{pe} \neq \mathbf{Z}$ and $\mathbf{Z}_{pe} < \mathbf{Z}$ in size because $\mathbf{Z}$ needs to be augmented for animals without observations, but that are included in the additive relationship matrix.

In model (13), the permanent environmental and residual effects for different animals are, in general, assumed to be uncorrelated. In addition, it is usually assumed that there is no correlation between additive and permanent environmental effects for the same animal, as well as between residual and permanent environmental effects. Therefore, the variances of both random effects and residual effects from model (13) can be written as follows:

$$\operatorname{var}\begin{bmatrix}\mathbf{a}\\ \mathbf{p}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{A}\sigma_a^2 & 0 & 0\\ 0 & \mathbf{I}\sigma_{pe}^2 & 0\\ 0 & 0 & \mathbf{I}\sigma_e^2\end{bmatrix} \qquad (14)$$

where σ_{pe}^2 is the variance due to permanent environmental effects; and the other terms were defined in (8).

Thus, the model illustrated in (13) is an extended simple animal model that is called repeatability animal model and its corresponding MME can be written as:

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{Z}_{pe} \\ \mathbf{Z}'\mathbf{X} & \mathbf{Z}'\mathbf{Z}+\mathbf{A}^{-1}\lambda_a & \mathbf{Z}'\mathbf{Z}_{pe} \\ \mathbf{Z}'_{pe}\mathbf{X} & \mathbf{Z}'_{pe}\mathbf{Z} & \mathbf{Z}'_{pe}\mathbf{Z}_{pe}+\mathbf{I}\lambda_{pe} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{Z}'_{pe}\mathbf{y} \end{bmatrix} \quad (15)$$

with λ_a as defined in (9) and $\lambda_{pe} = \frac{\sigma_e^2}{\sigma_{pe}^2}$.

The vector $\hat{\mathbf{p}}$ in (15) contains solutions for the permanent environmental effects. For an animal, this solution represents the environmental influences but also, if not accounted for in the model, the non-additive effects that are proper to the animal and affect its performance for life.

The basic assumption of a repeatability model is that repeated observations were regarded as expression of the same trait over time. A genetic correlation of 1 is assumed between repeated observations. Thus, it should be made clear that in any situation with multiple phenotypic observations for the same animal, it must be decided whether these observations are repeated measurements over time or are measurements on different, possibly correlated traits (*i.e.* with a genetic correlation between observations less than 1). In the first case, a repeatability model, while in the second case, a multiple-trait model ideally should be preferred.

ANIMAL MODEL WITH MULTIPLE GENETIC RANDOM EFFECTS

The above models and MME can be extended to account for multiple genetic random effects. Additional genetic random effects such as maternal effects, associative effects, non-additive effects can be fitted in the previous models. In the context of this thesis, only models with maternal effects are detailed below and models with non-additive effects are briefly introduce thereafter. Of course, there are other potential models fitting multiple genetic random effects.

MATERNAL ANIMAL MODEL

As explained previously, the performance of an individual depends on its additive genetic value and a random environmental deviation (permanent environmental effects and/or residual effects); however its performance may also be indirectly influenced by its mother other than through the genes she transmits on it (Willham, 1972). These indirect influences on the performance of an individual are called maternal effects and can be considered as the ability of dams to provide, in a broad sense, a suitable environment to their progenies. Maternal effects are strictly environmental for the progeny, but can have both genetic and environmental components. Similar to the genetic component of an individual, the maternal genetic component can be divided into additive and non-additive (dominance, epistasis) effects (Willham, 1963). It is the maternal additive genetic component of the dam that is passed on to all of her offspring, but it is expressed only when the female offspring have progeny of their own. The environmental part may be partitioned into permanent and temporary environmental (especially in sheep and swine) components in some cases, such as multiple litters with multiple offspring by litter for the same dam (Bijma, 2006). Additionally, the environmental part may also include the maternal non-additive effects of the dam.

In summary, the dam may contribute in two ways to the performance of the offspring: first, through her direct additive effects passed to the progeny (*i.e.* sample half of the dam’s nuclear genes) and second, through her ability to provide a suitable environment (*i.e.* genes of the dam responsible for the maternal effects) (Willham, 1972). Consequently, in animal selection, and especially in dam lines, it is important to consider these maternal effects in models used for genetic evaluation of traits influenced by those effects such as calving ease in dairy and beef cattle, birth weight and early growth in beef cattle and swine, or lamb survival to weaning and weaning weight in sheep for instance. Furthermore, ignoring maternal effects can bias the estimation of direct genetic effects. Willham (1963) presented a general quantitative genetic model for traits influenced by these maternal effects in which the sum of direct effects due to the individual and maternal effects due to its mother contribute to the trait. This Willham's model is readily incorporated into a mixed-model framework, allowing the estimation of direct additive as well as maternal additive and environmental effects (Quaas and Pollak, 1981).

In matrix notation, the Willham's model for maternally influenced traits can be written as:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_d\mathbf{a} + \mathbf{Z}_m\mathbf{m} + \mathbf{W}\mathbf{p}_m + \mathbf{e} \qquad (16)$$

in which $\mathbf{a}$ and $\mathbf{m}$ are vectors of direct and maternal additive random effects, respectively. The vector $\mathbf{a}$ has the same definition than in (7). The vector $\mathbf{p}_m$ contains the maternal environmental random effects which may include permanent and temporary environmental effects, but also the maternal non-additive effects of the dam. The known incidence matrices $\mathbf{Z}_d$, $\mathbf{Z}_m$, and $\mathbf{W}$ relate phenotypic observations ($\mathbf{y}$) to direct additive, maternal additive and permanent environmental effects, respectively. The vectors $\mathbf{b}$ and $\mathbf{e}$ and the known incidence matrix $\mathbf{X}$ were described previously in (7). In the Willham's model (16), direct and maternal additive effects can be correlated within an animal whereas the environmental effect of its dam and its direct additive effect are usually assumed to be uncorrelated as well as with its maternal additive effect. Therefore, the variances of random and residual effects from model (16) can be written as follows:

$$\operatorname{var}\begin{bmatrix} \mathbf{a} \\ \mathbf{m} \\ \mathbf{p}_m \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\sigma_a^2 & \mathbf{A}\sigma_{am} & 0 & 0 \\ \mathbf{A}\sigma_{am} & \mathbf{A}\sigma_m^2 & 0 & 0 \\ 0 & 0 & \mathbf{I}\sigma_{Pm}^2 & 0 \\ 0 & 0 & 0 & \mathbf{I}\sigma_e^2 \end{bmatrix} \qquad (17)$$

where σ_a^2 and σ_m^2 are the direct and maternal additive variances, respectively; σ_{am} is the direct-maternal additive covariance; σ_{Pm}^2 is the variance due to the maternal environmental random effects; and the other terms were defined in (8).

The genetic covariance (σ_{am}) plays a key role in the interpretation of estimated genetic parameters and the prediction of response to selection. However, this covariance is suggested to be sensitive to estimation bias (Robinson, 1996; Eaglen *et al.*, 2012). A number of theories were proposed to explain the origin of this bias in estimates of direct-maternal genetic covariances. These theories range from ignored environmental dam-offspring covariances (Koerhuis and Thompson, 1997; Bijma, 2006) to unaccounted contemporary groups (Robinson, 1996; Dodenhoff *et al.*, 1999; Gutiérrez *et al.*, 2006). This bias could also be due to the impact of data structure as the direct additive component is estimated from a single observation per animal, whereas maternal effects depend on the number of progeny per dam, the number of dams with recorded performance, and the number of generation of recorded data (Clément *et al.*, 2001; Maniatis and Pollott, 2003).

The model illustrated in (16) is an extended animal model accounting for the maternal effects and that can be called maternal animal model. The MME for a maternal animal model can be written as:

$$\begin{bmatrix} \mathbf{X'X} & \mathbf{X'Z}_d & \mathbf{X'Z}_m & \mathbf{X'W} \\ \mathbf{Z}'_d\mathbf{X} & \mathbf{Z}'_d\mathbf{Z}_d + \mathbf{A}^{-1}\lambda_a & \mathbf{Z}'_d\mathbf{Z}_m + \mathbf{A}^{-1}\lambda_{am} & \mathbf{Z}'_d\mathbf{W} \\ \mathbf{Z}'_m\mathbf{X} & \mathbf{Z}'_m\mathbf{Z}_d + \mathbf{A}^{-1}\lambda_{am} & \mathbf{Z}'_m\mathbf{Z}_m + \mathbf{A}^{-1}\lambda_m & \mathbf{Z}'_m\mathbf{W} \\ \mathbf{W'X} & \mathbf{W'Z}_d & \mathbf{W'Z}_m & \mathbf{W'W} + \mathbf{I}\lambda_{Pm} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{m}} \\ \hat{\mathbf{p}}_m \end{bmatrix} = \begin{bmatrix} \mathbf{X'y} \\ \mathbf{Z}'_d\mathbf{y} \\ \mathbf{Z}'_m\mathbf{y} \\ \mathbf{W'y} \end{bmatrix} \quad (18)$$

with $\mathbf{G}_0 = \begin{bmatrix} \sigma_a^2 & \sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$; $\mathbf{G}_0^{-1} = \begin{bmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{bmatrix}$; $\begin{bmatrix} \lambda_a & \lambda_{am} \\ \lambda_{am} & \lambda_m \end{bmatrix} = \sigma_e^2 \begin{bmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{bmatrix}$ and $\lambda_{Pm} = \dfrac{\sigma_e^2}{\sigma_{Pm}^2}$.

NON-ADDITIVE ANIMAL MODEL

Only purely additive gene action has been considered in the animal models presented above. However, these models can be extended to include non-additive gene action by adding an additional genetic random effect. Non-additive genetic effects are the interactions among alleles both within and across gene loci (*i.e.* dominance and epistasis effects respectively). These non-additive effects tend to be highly confounded with others, such as permanent environmental or maternal environmental effects. But when non-additive effects are significant, ignoring them in the animal model could make the estimation of additive effects less accurate, *e.g.* in the case of genetic evaluation of animals in a multi-breed population.

The MME for non-additive animal model can be set up in the same way as that for an additive animal model in the case of a maternal animal model. However, the addition of non-additive effects in model (7) requires the setting up of a non-additive relationship matrix (*i.e.* dominance or epistatic relationship matrix) which is more difficult to invert, especially in the case of large data sets. Though, methods for computing directly the inverse of a non-additive relationship matrix have been developed by Hoeschele and VanRaden (1991) for the dominance relationship matrix and by VanRaden and Hoeschele (1991) for the epistatic relationship matrix.

MULTIPLE-TRAIT ANIMAL MODEL

A multiple-trait animal model is an extension of the simple animal model but can also be considered as a case of model fitting multiple genetic random effects, as seen in the previous sections, just applied to different groups of records (traits). In a multiple-trait model, several traits can be analysed simultaneously taking correlation (if exists) between traits into account. Therefore, multiple-trait models can provide more accuracy as all information from correlated traits is used (van der Werf, 2002).

For ease of presentation, assume two correlated traits are available for simultaneous analysis under animal model. When animals are ordered within traits, the bivariate animal model for the two traits in matrix notation is:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix}=\begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix}+\begin{bmatrix}\mathbf{Z}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_2\end{bmatrix}\begin{bmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{bmatrix}+\begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix} \qquad (19)$$

in which for the *i*th trait (*i*=1 or 2), $\mathbf{y}_i$ is the vector of phenotypic observations; $\mathbf{b}_i$ is the vector of unknown fixed effects; $\mathbf{a}_i$ is the vector of unknown animal's additive random effects; $\mathbf{e}_i$ is the vector of the unknown residual effects, and $\mathbf{X}_i$ and $\mathbf{Z}_i$ are the known incidence matrices relating observations to corresponding fixed and random additive effects, respectively.

It is assumed that the covariance matrices for the additive effects and for the residual effects are defined, respectively, as:

$$\mathrm{var}\begin{bmatrix}\mathbf{a_1}\\ \mathbf{a_2}\\ \mathbf{e_1}\\ \mathbf{e_2}\end{bmatrix}=\begin{bmatrix}g_{11}\mathbf{A} & g_{12}\mathbf{A} & 0 & 0\\ g_{21}\mathbf{A} & g_{22}\mathbf{A} & 0 & 0\\ 0 & 0 & r_{11}\mathbf{I} & r_{12}\mathbf{I}\\ 0 & 0 & r_{21}\mathbf{I} & r_{22}\mathbf{I}\end{bmatrix}=\begin{bmatrix}\mathbf{G}_0\otimes\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_0\otimes\mathbf{I}\end{bmatrix}=\begin{bmatrix}\mathbf{G} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}\end{bmatrix} \qquad (20)$$

where $\mathbf{G}_0$ is the covariance matrix for the additive effects with each element defined as: g_{11} = additive variance for direct effects for trait 1; $g_{12} = g_{21}$ = additive covariance between both traits; $g_{22} =$ additive variance for direct effects for trait 2; $\mathbf{A}$ is the numerator relationship matrix among animals as in (8); $\otimes$ is the Kronecker product; $\mathbf{R}_0$ is the matrix of residual covariances with each element defined as: $r_{11} =$ variance for residual effects for trait 1; $r_{12} = r_{21} =$ residual covariance between both traits; $r_{22} =$ variance for residual effects for trait 2; and $\mathbf{I}$ is an identity matrix. As previously mentioned, residuals for a given trait measured on different individuals are, in general, assumed to be uncorrelated; this is not necessarily the case for different traits measured on the same individual. Indeed, the phenotypic correlation between traits is often the result of correlation between genetic but also environmental effects. It is also important to notice that the matrix of residual covariances shown in (20) is applicable when there are no missing observations for any trait. If not, the inverse of the residual covariance matrix $\mathbf{R}$ becomes a bit trickier.

Finally, the MME for the multiple-trait animal model described in (19) may be written as:

$$\begin{bmatrix} \mathbf{X}_1' r^{11}\mathbf{X}_1 & \mathbf{X}_1' r^{12}\mathbf{X}_2 & \mathbf{X}_1' r^{11}\mathbf{Z}_1 & \mathbf{X}_1' r^{12}\mathbf{Z}_2 \\ \mathbf{X}_2' r^{21}\mathbf{X}_1 & \mathbf{X}_2' r^{22}\mathbf{X}_2 & \mathbf{X}_2' r^{21}\mathbf{Z}_1 & \mathbf{X}_2' r^{22}\mathbf{Z}_2 \\ \mathbf{Z}_1' r^{11}\mathbf{X}_1 & \mathbf{Z}_1' r^{12}\mathbf{X}_2 & \mathbf{Z}_1' r^{11}\mathbf{Z}_1 + g^{11}\mathbf{A}^{-1} & \mathbf{Z}_1' r^{12}\mathbf{Z}_2 + g^{12}\mathbf{A}^{-1} \\ \mathbf{Z}_2' r^{21}\mathbf{X}_1 & \mathbf{Z}_2' r^{22}\mathbf{X}_2 & \mathbf{Z}_2' r^{21}\mathbf{Z}_1 + g^{22}\mathbf{A}^{-1} & \mathbf{Z}_2' r^{22}\mathbf{Z}_2 + g^{22}\mathbf{A}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}}_1 \\ \hat{\mathbf{b}}_2 \\ \hat{\mathbf{a}}_1 \\ \hat{\mathbf{a}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1'\left(r^{11}\mathbf{y}_1 + r^{12}\mathbf{y}_2\right) \\ \mathbf{X}_2'\left(r^{21}\mathbf{y}_1 + r^{22}\mathbf{y}_2\right) \\ \mathbf{Z}_1'\left(r^{11}\mathbf{y}_1 + r^{12}\mathbf{y}_2\right) \\ \mathbf{Z}_2'\left(r^{21}\mathbf{y}_1 + r^{22}\mathbf{y}_2\right) \end{bmatrix} \quad (21)$$

where g^{ij} and r^{ij} are elements of $\mathbf{G}_0^{-1}$ and $\mathbf{R}_0^{-1}$, respectively. In (21), it can be seen that both genetic and environmental correlations between traits are used simultaneously to take selection bias into account. If r^{12}, r^{21}, g^{12} and g^{21} are set to zero, the bivariate MME reduces to a simple combination of two single-trait analyses since the two traits become uncorrelated and there is no flow of information from one trait to the other.

In the bivariate animal model (19), the two traits are assumed to have the same model in common. However, the two traits might be affected by different fixed or random effects and so should be fitted with different models. Some animals may be represented in $\mathbf{y}_1$ but not in $\mathbf{y}_2$, or vice-versa. The multiple-trait animal model in (19) can be set up to analyse more than two correlated traits and becomes a multivariate animal model leading to more complex matrices $\mathbf{G}$ and $\mathbf{R}$, and then more complex MME. A common application of multiple-trait animal model is in evaluation of linear type traits. Moreover, one of the analysed traits can be a maternally influenced trait, making covariance structure and MME further complicated.

THE NON-LINEAR MIXED MODEL

In animal production, usually traits are expressed on continuous or on discrete scales. Traits, such as milk yield, body weight, height, are generally expressed and recorded on a continuous scale and are assumed to be normally distributed. Discrete scales can be ordered or not, using binary or more than two categories and therefore often called categorical scales. Not ordered discrete scales can be transformed to multiple-trait systems (yes-no), but are exceptional. Usual traits such as calving ease, survival, disease susceptibility, type scores are examples of traits expressed and recorded in ordered discrete categories. Some ordered discrete traits (*e.g.* number of piglets born) are counting products or similar items and therefore special cases of continuous distributions (*i.e.* two piglets is equal to exactly two times one piglet); others however do not display any apparently continuous phenotypic distribution (*e.g.* a more difficult calving is not two times more difficult than an easy one). These traits are, in general, called categorical or threshold traits.

Theoretically, statistical genetic analyses based on linear mixed models – as those presented in the previous sections – are generally unsuitable for discrete traits since most of the assumptions required are violated (Thompson, 1979; Gianola, 1982).

In the early 1980s, non-linear models were developed based on Wright's threshold concept for analysis of categorical traits in animal breeding (Gianola, 1982; Gianola and Foulley, 1983; Harville and Mee, 1984; Gilmour *et al.*, 1985). Models based on the threshold concept account for the categorical nature of the trait by assuming an underlying continuous variable called liability with truncation points – *i.e.* the thresholds – that determine the categories into which the phenotypic observations may fall (Wright, 1934; Gianola, 1982). With these models, it is possible to scale categorical phenotypic observations so as to conform to intervals of the continuous distribution and then applying linear models on the scaled data.

The most commonly known non-linear models based on the threshold concept and used to analyse categorical traits in animal breeding are the generalized mixed linear models using link function and the (ordinal) threshold models.

GENERALIZED LINEAR MIXED MODELS

Generalized linear mixed models (or GLMMs) can be considered as an extension of linear mixed models that can be used to analyse traits with a non-normal distribution. Moreover, GLMMs are models in which the independent variables (*i.e.* fixed and random effects) explain, indirectly through a function, the mean of the dependent variable. This is in contrast to classical linear mixed models in which the independent variables explain the dependent variable or its mean directly.

The GLMMs can be written in matrix notation as:

$$g(\boldsymbol{\mu}) = \mathbf{Xb} + \mathbf{Zu} \qquad (22)$$

where $\boldsymbol{\mu} = E(\mathbf{y})$ is the expectation or mean of the dependent variable y and $g(.)$ is the link function allowing the mean of the dependent variable y to be related to the linear combination of the fixed and random effects excluding the residuals. The choice of link function depends upon the distribution of the dependent variable. The most common link functions used in animal breeding are the logistic (logit model), Gaussian (probit model) as well as the Poisson regression (Poisson model).

From (22), the mean is:

$$\boldsymbol{\mu} = g^{-1}(\mathbf{Xb}+\mathbf{Zu}) \qquad (23)$$

where $g^{-1}(.)$ is the inverse link function, which is a function that transforms $\mathbf{Xb}+\mathbf{Zu}$ back to the mean. Thus, observations of the categorical variable y can be expressed as:

$$\mathbf{y} = g^{-1}(\mathbf{Xb}+\mathbf{Zu})+\mathbf{e} \qquad (24)$$

where $\mathbf{e}$ is a vector of random errors that can have a distribution other than normal.

THE (ORDINAL) THRESHOLD MODEL

Threshold models are strongly related to GLMMs, however the literature is not always very consistent in the way they are considered special cases of GLMMs or not.

The (ordinal) threshold model is appropriate for ordered categorical traits with more than 2 categories, as long as there are enough observations in each category. Again, the threshold model postulates a liability (L) for each phenotypic observation, such that the observed categorical variable takes value j if $t_{j-1} \leq L < t_j$, where t_{j-1} and t_j are thresholds, and $j = 1, 2, ..., j$ indexes the categories to which a phenotypic observation belongs (Gianola, 1982). The basic theory of (ordinal) threshold model with reference to animal breeding has been given by Gianola and Foulley (1983), and Harville and Mee (1984). In matrix notation, the (ordinal) threshold model can be written as:

$$\mathbf{L} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \qquad (25)$$

where $\mathbf{L}$ is a vector of unobserved liabilities on a normal scale. The conditional probability, given the effects, that observations fall in category *j* can be written as:

$$P(\mathbf{y}|\mathbf{b},\mathbf{u}) = \Phi\left[\mathbf{t}_j - (\mathbf{Xb} + \mathbf{Zu})\right] - \Phi\left[\mathbf{t}_{j-1} - (\mathbf{Xb} + \mathbf{Zu})\right] \quad (26)$$

where $\mathbf{y}$ is the vector of observed phenotypic observations of the categorical trait, $\Phi(.)$ is the normal cumulative distribution function and $\mathbf{t}_{j-1}$ and $\mathbf{t}_j$ are vectors with the t_{j-1} and t_j thresholds, respectively. The thresholds must satisfy $-\infty = t_0 \leq t_1 \leq ... \leq t_j = \infty$. Since L is not observed, it is not possible to solve for $\mathbf{u}$ using the usual MME (6). The solution to a threshold model is non-linear in computational complexity, and there must be back and forth calculations of thresholds and effects in the model until convergence of the system of equations stabilizes. Strategies for solving threshold MME have been proposed by Gianola and Foulley (1983) and Misztal *et al.* (1989).

APPLICATIONS OF MIXED MODEL METHODOLOGY

Theoretical aspects of most mixed models used for genetic evaluations in livestock production have been summarized in the previous sections. All of these models can be used as they are but can also be adapted in various ways depending upon the structure of the data, analysis final goal and according to the encountered problems. Hereafter, two examples of applications of mixed model methodology are provided and discussed.

TEST-DAY ANIMAL MODELS

Several traits of interest in animal production systems are measured repeatedly over time, for instance, bodyweight and yield traits (*e.g.* milk, eggs, wool).

In dairy cattle production systems, these repeated measurements over time (*e.g.* milk yield, fat and protein contents, somatic cell score) are typically referred to as test-day (**TD**) production records since they are specific to a particular testing day, such tests usually being distributed over the whole lactation. These measurements are collected by milk recording organizations around the world and they conventionally have been used in aggregated forms as lactation records in a lactation model, such as traditional 305-day approaches.

So as to address some problems encountered with the lactation models, statistical models analysing directly TD records have been introduced (Ptak and Schaeffer, 1993; Swalve, 1998). A test-day model (**TDM**) is defined as a statistical procedure which considers all genetic and environmental effects directly on TD basis allowing better modelling of factors affecting the performance of cow over the lactation and then possibly greater accuracy of genetic evaluations (Ptak and Schaeffer, 1993; Jamrozik *et al.*, 1997).

There are numerous advantages of using TDM compared with the traditional use of lactation models. The main advantage is its ability to account for time-dependent variation in the course of lactation (Swalve, 2000). It allows the environments specific to individual TD to be better taken into account, resulting in an improved accuracy of evaluations. Furthermore, the use of extended records for culled cows and for records in progress is not more required, and TDM is better suited to predict daily production, to detect outliers, thereby helping decision-making for management purposes (Mayeres *et al.*, 2004). There is also an economic advantage to use TDM due to its ability to include data from different recording schemes: some herds may only contribute milk yield while in others fat and protein contents are also sampled (Jamrozik *et al.*, 1997).

Based on Swalve (1998), TDM may generally be separated in two classes: The use of TD records in a one-step method directly providing EBV for dairy production, and two-step methods that apply some correction for environmental effects on the TD level but perform evaluations on records or residuals combined after the first step.

TWO-STEP TEST-DAY MODEL

In two-step TDM, the first step consists of a prediction step where the TD records are adjusted for TD environment such as, for example, herd TD effects, age-season effects, stage of lactation or milking frequency. These adjusted TD records are then combined into full lactation records weighting the individual TD record according to the correlations among them. The second step consists of a step of analysis using the current animal model procedures; thus it is an indirect use of TD records to produce EBV for dairy production (Swalve, 2000; Gengler, 2002).

A two-step TDM was performed in New Zealand from 1996 until 2007 in an across-breed genetic evaluation system for production traits of New Zealand dairy cattle (Harris *et al.*, 1996). Australia (*e.g.* Jones *et al.*, 1990) and North-eastern United States (*e.g.* Stanton *et al.*, 1992; Van Tassell *et al.*, 1992) used also two-step TDM in the past.

ONE-STEP TEST-DAY MODEL

One-step TDM makes a direct use of TD records by combining the first and the second steps of two-step TDM (Gengler, 2002). One-step TDM have been derived from repeatability animal models under which TD records within lactation are taken as repeated measurements on the same trait, and a permanent environmental effect accounts for environmental similarities between different TD within the same lactation. Two approaches in one-step TDM can be distinguished depending upon whether or not there are cow specific regression effects; *i.e.* whether or not the one-step TDM allows the modelling of the cow specific lactation shape functions by using random regressions (Swalve, 2000). Both of these approaches are discussed below.

Fixed regression test-day model

In the fixed regression approach, the lactation curve is considered as a fixed effect and the random component of the model are specified as a traditional repeatability animal model. The curvilinear pattern of lactation curve is fitted as a fixed effect by several coefficients of a fixed regression of yield on days in milk (**DIM**) or rather a function of these DIM (*e.g.* Wilmink's function, Legendre polynomial). These regressions are typically nested within subclasses of fixed effects to account for different lactation curve shapes for groups of animals which may be defined for instance, by years of birth, parity number, age and season of calving within parities (Ptak and Schaeffer, 1993). Thus, with a fixed regression TDM, the phenotypic shape of the lactation curve is assumed to be the same for all cows, being the height of the curve different from cow to cow.

As in a repeatability animal model, genetic variation and permanent environmental variation in the course of lactation are assumed be constant and hence the genetic and permanent environmental correlations between yields at different DIM are assumed constant and close to one regardless of time elapsing between TD records. However, several studies have demonstrated that correlation between TD records decrease as interval between TD increases (Ptak and Schaeffer, 1993; Swalve, 1995). Therefore, the assumption that the variances of random effects are homogenous throughout the lactation was difficult to justify.

Random regression test-day model

Random regression models are useful for modelling traits that change gradually and continually with time and are measured repeatedly on individuals, such as TD production records in dairy cattle. Accordingly, Schaeffer and Dekkers (1994) extended the fixed regression model for the analysis of TD records in dairy cattle by using random regression coefficients for each individual random factor (*i.e.* additive, permanent environmental effects and for any other random effects). In this approach, the lactation curve for an individual cow could be seen as two sets of regressions on DIM. Fixed regressions for all cows belonging to the same subclasses of fixed effects, as explained above, describe the phenotypic shape for that cow, and the random regressions for a cow describe the deviation from this phenotypic shape, allowing cows to have differently shaped lactation curves (Jamrozik *et al.*, 1997). Therefore, with a random regression TDM, each cow can have its own lactation curve, its own shape.

With random regression TDM, the shape of the lactation curve is assumed to be also influenced by genetic and permanent environmental random effects. Random regression TDM can accommodate changes in the genetic and permanent environmental covariance structure of TD records during the lactation by fitting functions of times (Jensen, 2001). Similarly, genetic and permanent environmental correlations between TD records at different DIM can be less than one. Moreover, as the genetic merit of an individual is allowed to change for any day in the lactation, random regression TDM offer the opportunity to express EBV as curves of genetic merit (Schaeffer and Dekkers, 1994; Swalve, 2000). Furthermore, the genetic variance and “genetic yields” for each single day of lactation can be estimated and used to define suitable criteria of persistency, *i.e.* a trait of economic importance due to its impact on feed costs, health, and fertility (Gengler, 1996).

The matrix notation of a random regression TDM used may be written as:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{QZa} + \mathbf{Q}_{pe}\mathbf{Z}_{pe}\mathbf{p} + \mathbf{e} \qquad (27)$$

where $\mathbf{y}$ is the vector of TD records; $\mathbf{b}$ is a vector of fixed effects, which are independent of the time scale for the TD records, and fixed regression coefficients; and $\mathbf{a}$ and $\mathbf{p}$ are here vectors of random regression coefficients for animal additive and permanent environmental effects, respectively; $\mathbf{e}$ is the vector of the unknown residual effects; and $\mathbf{X}$ and $\mathbf{Z}$ are the known incidence matrices. The matrices $\mathbf{Q}$ and $\mathbf{Q}_{pe}$ contains the covariates associated with DIM. Covariates are, in general, simple functions of DIM such as polynomials, orthogonal polynomials (*e.g.* Legendre polynomials), splines or parameters of lactation functions, such as

Wilmink's function (Wilmink, 1987) or Ali and Schaeffer's function (Ali and Schaeffer, 1987) for example. The most used functions are the Legendre polynomials because they make no assumption about the shape of curve and are easy to apply. In addition, they describe efficiently the evolution of milk yield during a complete lactation of dairy cows in different management conditions (Gengler *et al.*, 1999; Brotherstone *et al.*, 2000).

With the random regression TDM illustrated in (27), the covariance matrices for the random effects are assumed defined as:

$$\text{var}\begin{bmatrix}\mathbf{a}\\\mathbf{p}\\\mathbf{e}\end{bmatrix}=\begin{bmatrix}\mathbf{A}\otimes\mathbf{G}_0 & \mathbf{0} & \mathbf{0}\\\mathbf{0} & \mathbf{I}\otimes\mathbf{P}_0 & \mathbf{0}\\\mathbf{0} & \mathbf{0} & \mathbf{I}\otimes\mathbf{R}_0\end{bmatrix} \qquad (28)$$

in which $\mathbf{G}_0$ is here the covariance matrix of the additive genetic random regression coefficients; $\mathbf{P}_0$ is the covariance matrix of the permanent environmental random regression coefficients; $\mathbf{R}_0$ is a matrix of residual variances; $\mathbf{A}$ is the additive relationship matrix; $\otimes$ is the Kronecker product; and $\mathbf{I}$ is an identity matrix.

A random regression TDM used with a multiple-trait animal model was introduced for genetic evaluation of milk production traits in Canada by Schaeffer *et al.* (2000). Since then, random regression TDM have become the models of choice for genetic evaluation of production traits in dairy cattle. Several ways of the methodology have been used for genetic evaluation of production traits in dairy cattle. Most applications of random regression TDM have been multiple trait systems, in which multiple characters and multiple lactations are considered together (*e.g.* Auvray and Gengler, 2002; Muir *et al.*, 2007; Konstantinov *et al.*, 2009). Random regression TDM can also be applied in joint evaluations across countries system (*e.g.* Emmerling *et al.*, 2002; de Roos *et al.*, 2004), which adds further complexity to the applied random regression TDM.

However, even if random regression models are widely used in dairy cattle, they are also used in other areas of animal breeding (*e.g.* dairy sheep, egg production, growth traits), and they can be also extended to include random regressions for maternal genetic and maternal permanent environmental effects for instance. A review about the numerous applications of random regression models in animal breeding was provided by Schaeffer (2004).

GENETIC EVALUATION IN CROSSBREEDING: MULTI-BREED MODELS

Crossbreeding is the method of mating individuals from different lines, breeds or populations to produce crossbred offspring. Crossbreeding is applied to capitalise on breed (line) complementarity and heterosis. Breed (line) complementarity allows bringing together desirable characteristics from different pure breeds or lines that are difficult to select together because of genetic antagonisms (Bidanel, 1992). Heterosis, also known as hybrid vigour, is the increase in performance above what is expected based on the parents of the crossbred offspring. Heterosis is a non-additive effect due to dominance effects (*i.e.* interactions within loci) and epistasis effects (*i.e.* interaction between loci) and so its importance depends upon the genetic distance between the parental breeds (Bidanel, 1992; Swan and Kinghorn, 1992). Crossbreeding can also provide an opportunity to make progress in one generation that would require generations of selection to obtain it and can allow the introduction of a new breed in a herd (Bidanel, 1992).

Crossbreeding is broadly used in beef cattle, swine, and poultry production systems (*e.g.* Pollak and Quaas, 1998; Lutaaya *et al.*, 2002; Yang and Jiang, 2005; Ibánẽz-Escriche *et al.*, 2011). These species rely heavily on the heterosis and breed complementarity effects to improve productivity and efficiency of production in the crossbred offspring (Hansen, 2006). In dairy cattle, crossbreeding between *Bos taurus* dairy breeds and local *Bos indicus* cattle is common in tropical climates (Cunningham and Syrstad, 1987; McDowell *et al.*, 1996), where high-producing dairy breeds are less adapted to the environment than local breeds. In contrast, in temperate climates, crossbreeding in dairy cattle is not used extensively except in New Zealand (Sørensen *et al.*, 2008). Pure breeding systems in dairy cattle are indeed widespread in most temperate countries, largely due to the advantage of Holstein cattle in milk volume and the strong historical influence of purebred breeders and breed associations (Cunningham and Syrstad, 1987; Weigel and Barlass, 2003).

Nevertheless, for several years, dairy producers and breeders from most western countries have become increasingly interested in crossbreeding aspects for several reasons. First, the breeding goals of these countries have evolved in recent years, and animal selection is now based on economic indexes that do not include only production and conformation traits, but also consider functional traits such as female fertility, health, calving ease, and longevity (Miglior *et al.*, 2005; Nielsen *et al.*, 2006; Oltenacu and Broom; 2010). Although the functional traits are often still secondary to milk yield, they play an important role in increasing economic profit by reducing farm costs (McAllister, 2002). This evolution is mainly due to the increasing deterioration of the functional traits over the years within most major dairy breeds (*e.g.* Kadarmideen *et al.*, 2000; Heins *et al.*, 2006a; Sørensen *et al.*, 2008; Walsh *et al.*, 2011). It results from the high selection pressure on milk yield traits and genetic antagonisms between dairy production and functional traits, as well as the intensification of dairy production systems (Boettcher, 2005; Walsh *et al.*, 2011). Thus, this economic emphasis on functional traits combined with crossbreeding may be helpful to overcome these problems of decreasing functionality in dairy herds while maintaining a good milk production (Sørensen *et al.*, 2008). Several studies have indeed stated that crossbreeding had more influence on functional traits than on production (Heins *et al.*, 2006b; Blöttner *et al.*, 2011; Heins *et al.*, 2012). Second, inbreeding rates are growing rapidly in most of the high-producing dairy breeds due to efficient genetic selection programs (*e.g.* Sørensen *et al.*, 2005; Croquet *et al.*, 2006; Miglior *et al.*, 2008) and crossbreeding may be useful to reduce the impact of inbreeding depression by using all available genetic resources and taking advantage of the increase of heterozygosity (Weigel and Barlass, 2003; Hansen, 2006). Third, the trends in milk pricing favouring high fat and protein content have encouraged some milk producers to consider crossbreeding as a way to enhance milk nutrient contents, and this has enhanced the ability of other breeds and breed crosses to compete with Holsteins on an economic basis, especially in those countries where a great emphasis is given to the cheese industry (VanRaden and Sanders, 2003; Dal Zotto *et al.*, 2007). Finally, a greater income can be also obtained from selling crossbred calves as compared to purebred dairy calves, especially with beef x dairy crossbred calves that have good meat features (Wolfová *et al.*, 2007; Penasa *et al.*, 2009). A common European practice is to inseminate dairy and dual-purpose cows with semen derived from beef bulls if the calves will not be used for milk production in the dairy farms (Dal Zotto *et al*., 2009).

Most genetic evaluations carried out in livestock production compare animals only within breed, based on purebred information, so excluding information from crossbred animals. Often, in crossbreeding programs, purebred sires are selected among sires of a given breed after the breed has been selected for the program. The same applies to the selection of purebred dams. Therefore, it is assumed that the performance of crossbred progeny from the selected parents are predictable based on their purebred progeny and purebred relatives, *i.e.* assuming a correlation close to 1 between purebred and crossbred performance (Newman and Reverter, 2000). This is not the case for all livestock species and for all traits of economic importance. Indeed, genetic correlations less than 1 between purebred and crossbred performance were pointed out for economically important traits, for instance, by Newman *et al.* (2002) in beef, by Zumbach *et al.* (2006) in pigs or by Wei and van der Werf (1995) in poultry. Genetic correlation between purebred and crossbred performance decreases with increasing non-additive effects (particularly dominance) or gene frequency difference between parental populations (Wei *et al.*, 1991). Besides, environmental differences may exist between purebred and crossbred animals. For example, in commercial swine production, crossbreds are usually raised in environments of lower quality than those of purebreds concerning management, hygiene status and space per pig. Therefore, differences between purebred and crossbred performance may also reflect genotype by environment interactions (Bryant *et al.*, 2007; Dufrasne *et al.*, 2011). Thus, performance of purebred parents can be a poor predictor of performance of their crossbred progeny due to genetic and environmental differences between purebreds and crossbreds (Ibáñẽz-Escriche *et al.*, 2011).

Since the final goal of a crossbreeding program usually focuses on an increase of the crossbred performance, there is an interest to evaluate crossbred performance jointly with purebred performance to perform combined crossbred-purebred selection (Wei and van der Werf, 1994). To this end, linear mixed models used for genetic evaluation within breeds have been extended to accommodate information from animals of different breeds (purebred animals) but also information from crossbred animals, thereby enabling a joint evaluation of purebreds (from different breeds/lines) and crossbreds as well as an across-breed selection and mating programs. In the scientific literature, these types of models are mainly termed as “multi-breed models” but they may be referred to as “across-breed models”, “crossbred models” or even “combined crossbred-purebred models”. Hereafter, the term “multi-breed model” will only be used throughout this section.

Multi-breed models allow evaluating purebred animals more accurately for performance in crossbreeding compared to strictly single-breed or within-breed models due to the use all available progeny information (Lutaaya *et al.*, 2001; VanRaden *et al.*, 2007). Consequently, a re-ranking of purebred candidates to selection as well as a greater response to selection may occur when a combined crossbred-purebred evaluation is performed (Wei and van der Werf, 1994; Ibánẽz-Escriche *et al.*, 2011). Another benefit of the multi-breed models is that breed differences and heterosis are routinely estimated and updated (VanRaden *et al.*, 2007). Naturally, these models rely heavily on accurate recording of phenotypic observations and pedigrees of crossbred animals, which can be difficult to collect routinely in some cases (*e.g.* crossbred animals reared under industrial environment) and so could increase the investment in the selection programs (Dekkers, 2007). Furthermore, such models hold for only a limited number of generations and genetic parameters, such as correlation between purebred and crossbred performance, in the model should be frequently updated (Wei and van der Werf, 1994).

Multi-breed genetic evaluation models permit direct comparison of animals of various breed compositions and therefore allow genetic improvement in multi-breed populations. The multi-breed models have been mainly developed for genetic evaluations in swine and beef cattle populations where the use of crossbreeding is most widespread. Most of these multi-breed models involve the simultaneous adjustment of fixed effects and the estimation of breed differences and breed interactions (*e.g.* heterosis, dominance, recombination loss), as well as animals’ EBV. The inclusion of those breed effects and interactions (especially heterosis effects) prevents biased estimates of genetic parameters and therefore, avoids a bias in the estimation of EBV (Van Der Werf and De Boer, 1989).

In the following sub-sections, multi-breed models are presented into two main approaches: the single-trait approach, in which observations on purebred and crossbred animals are regarded as coming from the same trait, and the multiple-trait approach, in which observations on purebred and crossbred animals are considered as separate but correlated traits.

SINGLE-TRAIT APPROACH

Several methods are available to model breed additive and breed interactions mean effects for performance of crossbred animals in a multi-breed model (Van Vleck, 1997) using a single-trait approach. For instance, genetic groups provide an excellent mechanism to incorporate breed mean effects into a multi-breed model. Thus, unknown parent groups defined separately by breed or by breed composition can be fitted in animal model evaluations taking into account genetic means across breeds' changes (*e.g.* Arnold *et al.*, 1992; VanRaden *et al.*, 2007; Sánchez *et al.*, 2008). These groups can also be used for fitting the heterosis mean effects by defining groups separately by breed combinations (*e.g.* Sánchez *et al.*, 2008). Models with genetic groups have been discussed earlier in a previous section. Regressions on breed composition and regressions on heterozygosity fractions can be also a good alternative for fitting breed and heterotic mean effects, respectively, in a multi-breed model (*e.g.* Pollak and Quaas, 1998; Legarra *et al.*, 2007). Likewise, breed and heterotic mean effects can be accounted for by using estimates from the scientific literature to pre-adjust records, providing that the published estimates are reliable and applicable to the population being evaluated (*e.g.* Sullivan *et al.*, 1999).

To estimate accurately animals' genetic merit, it is important to properly model additive covariances between relatives to use most efficiently performance recorded on relatives in genetic evaluation. The covariance theory for a purebred population cannot be applied in the case of a multi-breed population. Since a multi-breed population is composed of several breed groups, each of which can be formed by purebred animals or by crossbred animals, the assumption that additive (co)variances are equal across breeds groups is not really appropriate although some studies made this assumption (Pollak and Quaas, 1998; Sullivan *et al.*, 1999; Legarra *et al.*, 2007; Sánchez *et al.*, 2008). Due to the heterogeneity of additive (co)variances across breed groups, the additive covariance matrix ($\mathbf{G}$) cannot be expressed as $\mathbf{A}\sigma_a^2$ as it is usually assumed in the simple animal model. Therefore, the useful method developed by Henderson (1976) and Quaas (1976) to construct directly the inverse of the relationship matrix ($\mathbf{A}^{-1}$) cannot be used for a multi-breed population. To deal with this, variance adjustment factors can be used on phenotypic data to adjust variances separately by breed to make genetic variance equal to that from a breed or a population of interest. Thus, the Henderson's method can be used to construct directly $\mathbf{A}^{-1}$ to perform genetic evaluation as in a conventional within-breed genetic evaluation. For instance, this adjustment method is currently used in the genetic evaluations of US mixed-breed dairy cattle populations to make genetic variance equal to Holstein base cows (VanRaden *et al.*, 2007).

However, procedures to compute directly $\mathbf{G}$ and $\mathbf{G}^{-1}$ with heterogeneous variances across breed groups for multi-breed populations have been proposed by Elzo (1986). He computed the additive variance for a crossbred group as a weighted mean of the additive variances of the parental pure breeds plus one half of the covariance between parents, where each weight is the proportion of the corresponding pure breed in the crossbred. This procedure was used by Arnold *et al.* (1992) to account for heterogeneity of variances among breed groups in a multi-breed beef cattle population. In contrast, Lo *et al.* (1993) showed that this method did not always lead to the correct additive variance for crossbred animals because it did not account for segregation variances. The segregation variance is the amount by which the additive variance in the second generation (F2) breed group exceeds that in the first generation (F1) breed group and is due to the differences in allelic frequencies between the parental pure breeds (Lande, 1981). Therefore, Lo *et al.* (1993) proposed an approach to incorporate segregation variance in the construction of $\mathbf{G}$ for a multi-breed population comprising an arbitrary number of pure breeds and all crosses involving these breeds, under a model with additive inheritance and multiple unlinked loci. According to their approach, the additive variance of animal i in a multi-breed population under the previous conditions can be computed as:

$$Var(a_i) = \sum_{p=1}^{P} f_p^i \sigma_p^2 + \frac{1}{2} Cov(a_j, a_k) + 2\sum_{p=1}^{P} \sum_{p'>p}^{P} \left(f_p^j f_{p'}^j + f_p^k f_{p'}^k \right) \sigma_{pp'}^2 \quad (29)$$

where P is the number of breeds involved in the founder generation, f_p^i is the proportion of genes of the animal *i* coming from breed *p*, *j* and *k* are the sire and the dam of animal *i*, $f_p^j f_{p'}^j$ ($f_p^k f_{p'}^k$) are the proportion of genes of the sire (dam) coming from breed *p* and *p'*, respectively, σ_p^2 is the additive variance of breed *p*, and $\sigma_{pp'}^2$ is the segregation variance between breeds *p* and *p'*. Except for terms involving the segregation variances, expression (29) is equivalent to the expression given by Elzo (1986) for the additive variance of a crossbred individual. Thus, Elzo's formula gives the same value of additive variance for a crossbred individual than (29) only when both parents are purebreds since the segregations terms are equal to zero for F_1 crossbreds. If additive variance of a crossbred animal is computed as described in (29), the same rules as for purebred populations (Henderson, 1976) can be used to obtain covariances between crossbred relatives. So, the additive covariance between two crossbred related animals *i* and *i'* can be computed as:

$$Cov\left(a_i, a_{i'}\right) = \frac{1}{2}\left[Cov\left(a_j, a_{i'}\right) + Cov\left(a_k, a_{i'}\right)\right] \quad (30)$$

which is the average covariance between additive values for animal *i'* and the parents *j* and *k* of animal *i*. Lo *et al.* (1993) provided rules to build a single additive covariance matrix $\mathbf{G}$ in which each animal is allowed to have unique breed composition and each breed has individual genetic variance. Furthermore, these authors also proposed an efficient method to calculate directly $\mathbf{G}^{-1}$, closely derived from the conventional rules to obtain $\mathbf{A}^{-1}$. This particular definition of $\mathbf{G}$ allows a multi-breed population to be analysed as a conventional animal model (7) and then can be used in conventional animal MME.

Nonetheless, due to the blending of breed-specific and segregation variances into $\mathbf{G}$ (29), some difficulties can be encountered to estimate the genetic (co)variance components for the multi-breed animal models (Birchmeier *et al.*, 2002; Cardoso and Tempelman, 2004) and when (co)variance components are estimated, its use in practice can be numerically expensive. In response to this problem, García-Cortés and Toro (2006) proposed to decompose $\mathbf{G}$ by source of variability as following:

$$\mathbf{G} = \sum_{p} \mathbf{A}_p \sigma_p^2 + \sum_{p} \sum_{p>p'} \mathbf{A}_{pp'} \sigma_{pp'}^2 \qquad (31)$$

where $\mathbf{A}_p$ and $\mathbf{A}_{pp'}$ are the partial additive relationship matrices due to pure breeds and segregation terms, respectively; σ_p^2 is the additive variance of breed *p*, and $\sigma_{pp'}^2$ is the segregation variance between breeds *p* and *p'*. The covariance structure of additive values (*i.e.* the relationships matrices) is built externally as in a conventional animal model (8), thus supplying a much simpler formulation for genetic variance components estimation, which is easy to assimilate with the estimation techniques available in general purpose software.

As a result, a multi-breed animal model using this specific decomposition of $\mathbf{G}$ can be written as:

$$\mathbf{y} = \mathbf{Xb} + \sum_{p} \mathbf{Za}_p + \sum_{p} \sum_{p>p'} \mathbf{Za}_{pp'} + \mathbf{e} \qquad (32)$$

where $\mathbf{a}_p$ and $\mathbf{a}_{pp'}$ are vectors of additive values (*i.e.* EBV) split by origin of variability. The vectors $\mathbf{y}$, $\mathbf{b}$, and $\mathbf{e}$ and the known incidence matrices $\mathbf{X}$ and $\mathbf{Z}$ were described in (7). The multi-breed animal model displayed in (32) is, therefore, a conventional animal model with multiple genetic random effects such as that described in “Maternal animal model” section but there is no covariance between these genetic effects in (32).

Assuming a case with two pure breeds (breeds A and B), the model (32) becomes:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za}_A + \mathbf{Za}_B + \mathbf{Za}_{AB} + \mathbf{e} \qquad (33)$$

in which $\mathbf{a}_A$, $\mathbf{a}_B$ and $\mathbf{a}_{AB}$ are the additive values split by origin of variability. The corresponding MME can be written as:

$$\begin{bmatrix} \mathbf{X'X} & \mathbf{X'Z} & \mathbf{X'Z} & \mathbf{X'Z} \\ \mathbf{Z'X} & \mathbf{Z'Z} + \mathbf{A}_A^{-1}\lambda_A & \mathbf{Z'Z} & \mathbf{Z}_d'\mathbf{Z} \\ \mathbf{Z'X} & \mathbf{Z'Z} & \mathbf{Z'Z} + \mathbf{A}_B^{-1}\lambda_B & \mathbf{Z}_m'\mathbf{Z} \\ \mathbf{Z'X} & \mathbf{Z'Z} & \mathbf{Z'Z} & \mathbf{Z'Z} + \mathbf{A}_{AB}^{-1}\lambda_{AB} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}}_A \\ \hat{\mathbf{a}}_B \\ \hat{\mathbf{a}}_{AB} \end{bmatrix} = \begin{bmatrix} \mathbf{X'y} \\ \mathbf{Z'y} \\ \mathbf{Z'y} \\ \mathbf{Z'y} \end{bmatrix} \qquad (34)$$

with $\lambda_x = \dfrac{\sigma_e^2}{\sigma_x^2}$

As illustrated in (34), each single EBV ($\hat{\mathbf{a}}$) is split into its breed-specific components ($\hat{\mathbf{a}}_A$ and $\hat{\mathbf{a}}_B$) and breed segregation component ($\hat{\mathbf{a}}_{AB}$) being each of these components random effects with specific relationship matrices $\mathbf{A}_A$, $\mathbf{A}_B$ and $\mathbf{A}_{AB}$, respectively. These components are assumed to be uncorrelated.

The main difference between the model proposed by García-Cortés and Toro (2006) and other models involving, for instance, permanent environmental effects (14), is the particular nature of the partial relationship matrices. In fact, the complete additive values in the conventional animal model are replaced by the partial additive values (32-34). The required inverses of these partial additive relationship matrices in (34) can be directly obtained by slightly changing the conventional method to compute $\mathbf{A}^{-1}$ (Quaas, 1976). The adapted algorithms to construct directly the inverse of the partial relationship matrices were provided by García-Cortés and Toro (2006).

The equivalence of this model (32) with the additive one proposed by Lo et *al.* (1993) for evaluating multi-breed populations has been shown empirically by García-Cortés and Toro (2006) and also by Munilla Leguizamon and Cantet (2010), which they provided the formal derivation using a Bayesian approach. In addition, the latter showed how to extend the model (32) to include maternal effects.

MULTIPLE-TRAIT APPROACH

As genetic correlations between purebred and crossbred performance are lower than 1 for several economically important traits, a straightforward way to combine performance of purebred animals with crossbred relatives information could be the use of multiple-trait model as seen previously, in which purebred and crossbred performance are treated as different traits with genetic correlation between them (Swan and Kinghorn, 1992).

For a two-breed terminal crossbreeding system, Wei and van der Werf (1994) described a combined crossbred and purebred selection method using a multiple-trait approach *via* reduced animal model. This model, which contains two additive effects, considers performance of purebred animals from breeds A and B and those from their terminal crosses AB as three different traits with genetic correlation between purebred and crossbred animals. Thus, the model proposed by Wei and van der Werf (1994) can be written in matrix notation as:

$$\begin{bmatrix} \mathbf{y}_A \\ \mathbf{y}_B \\ \mathbf{y}_{AB} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_A & 0 & 0 \\ 0 & \mathbf{X}_B & 0 \\ 0 & 0 & \mathbf{X}_{AB} \end{bmatrix} \begin{bmatrix} \mathbf{b}_A \\ \mathbf{b}_B \\ \mathbf{b}_{AB} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_A & 0 & 0 & 0 \\ 0 & 0 & \mathbf{Z}_B & 0 \\ 0 & \mathbf{Z}_{A,AB} & 0 & \mathbf{Z}_{B,AB} \end{bmatrix} \begin{bmatrix} \mathbf{a}_A \\ \mathbf{a}_{A,AB} \\ \mathbf{a}_B \\ \mathbf{a}_{B,AB} \end{bmatrix} + \begin{bmatrix} \mathbf{e}_A \\ \mathbf{e}_B \\ \mathbf{e}_{AB} \end{bmatrix} \quad (35)$$

where $\mathbf{y}_A$ and $\mathbf{y}_B$ are vectors of phenotypic observations for purebreds from breed A and B, respectively; $\mathbf{y}_{AB}$ is the vector of phenotypic observations for crossbreds; $\mathbf{b}_A$, $\mathbf{b}_B$ and $\mathbf{b}_{AB}$ are vectors of unknown fixed effects; $\mathbf{a}_A$ and $\mathbf{a}_B$ are vectors of additive effects of animals in breed A and B, respectively; $\mathbf{a}_{A,AB}$ and $\mathbf{a}_{B,AB}$ are vectors of additive effects of animals originating from breed A and B, respectively, for crossbred performance; $\mathbf{e}_A$, $\mathbf{e}_B$ and $\mathbf{e}_{AB}$ are vectors of residual effects; $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices relating observations to corresponding effects. The use of separate fixed effects enables capturing the general level of heterosis in crossbreds. Under a reduced animal model, performance of crossbred animals

$\left(\mathbf{y}_{AB}\right)$ are written in (35) as function of their parental additive effects for crossbred performance (*i.e.* $\mathbf{a}_{A,AB}$ and $\mathbf{a}_{B,AB}$) because the additive values of crossbreds are not of interest, since not used for breeding in the case of terminal crossbreeding system.

This multi-breed model (35) provides two EBV for each purebred animal: one related to mating within breed (*i.e.* $\mathbf{a}_A$ or $\mathbf{a}_B$) and another related to mating to another breed to produce the cross (*i.e.* $\mathbf{a}_{A,AB}$ .or $\mathbf{a}_{B,AB}$). Both these EBV are different but correlated by the relationship between purebred and crossbred performance. The model also assumes different variances in the two pure breeds and in the crossbreds, which is often the case in practice, and it also assumes different covariances among half-sib groups, dependent of the breed of the common parent. Furthermore, this model allows for low and breed-specific genetic correlation between the additive effects of purebreds and crossbreds. Hence, the covariance matrix for the additive effects ($\mathbf{G}$) in (35) is assumed to be:

$$\mathbf{G} = \mathrm{var}\begin{bmatrix} \mathbf{a}_A \\ \mathbf{a}_{A,AB} \\ \mathbf{a}_B \\ \mathbf{a}_{B,AB} \end{bmatrix} = \begin{bmatrix} \sigma^2_{AA}\mathbf{A}_A & \sigma_{A,AB}\mathbf{A}_A & 0 & 0 \\ \sigma_{A,AB}\mathbf{A}_A & \sigma^2_{AB/A}\mathbf{A}_A & 0 & 0 \\ 0 & 0 & \sigma^2_{BB}\mathbf{A}_B & \sigma_{B,AB}\mathbf{A}_B \\ 0 & 0 & \sigma_{B,AB}\mathbf{A}_B & \sigma^2_{AB/B}\mathbf{A}_B \end{bmatrix} = \begin{bmatrix} \mathbf{G}_A \otimes \mathbf{A}_A & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_B \otimes \mathbf{A}_B \end{bmatrix} \quad (3$$

where σ^2_{AA} and σ^2_{BB} are the additive variances of breed A and B, respectively; $\sigma^2_{AB/A}$ and $\sigma^2_{AB/B}$ are additive variance of breed A and B alleles in crossbreds (AB), respectively; $\sigma_{A,AB}$ and $\sigma_{B,AB}$ are additive covariances between purebred parents in breed A and B and their crossbred (AB) progeny; $\mathbf{A}_A$ and $\mathbf{A}_B$ are additive relationship matrices for breed A and B, respectively; and $\mathbf{G}_A$ and $\mathbf{G}_B$ are the additive covariance matrices of the purebred animals for breed A and B, respectively.

The covariance matrix for residuals ($\mathbf{R}$) in (35) is assumed to be a diagonal matrix:

$$\mathbf{R} = \mathrm{var}\begin{bmatrix} \mathbf{e}_A \\ \mathbf{e}_B \\ \mathbf{e}_{AB} \end{bmatrix} = \begin{bmatrix} \sigma^2_{eA}\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma^2_{eB}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma^2_{eAB}\mathbf{I} \end{bmatrix} \qquad (37)$$

where σ^2_{eA}, σ^2_{eB} and σ^2_{eAB} are the residual variances for the three breed groups (A, B or AB), respectively. Residual effects are uncorrelated because an individual can only belong to one of these three groups.

Since crossbreeding also allows exploiting non-additive variability, like dominance for instance, and that genetic correlations between purebred and crossbred performance are also influenced by dominance level, Lo *et al.* (1997) have extended the multi-breed model proposed by Wei and van der Werf (1994) to account for dominance covariances. Adding dominance effects into the model (35), it further adds one equation for each individual in breed group A, one for each individual in B, and one for each full-sib family in AB (Lo *et al.*, 1997). Their model accounts for all additive and dominance covariances when crossbreds are limited to F_1, which is the case in a two-breed terminal crossbreeding system. Therefore, this model is much simpler in comparison to the model for all types of crossbreds. In fact, theory involving covariances of multi-way crosses under dominance is very complicated because, in this situation, 25 parameters are needed to model these genetic covariances between purebred and crossbred animals, and the model complexity increases as more breeds are involved in the crossbreeding system (Lo *et al.*, 1995). While the number of required parameters is reduced to 9 to model genetic covariances (6 additive and 3 non-additive parameters) with the multi-breed model proposed by Lo *et al.* (1997) for a two-breed terminal crossbreeding system with no inbreeding.

In comparison with the use of single-breed model, the use of a multi-breed model as defined by Lo *et al.* (1997) is very worthwhile when the interest is in both purebred and crossbred evaluations (*i.e.* when genetic correlations between purebred and crossbred performance are low), and when the amount of crossbred information relative to purebred information is substantial. This model is also interesting when some traits are recorded on purebreds and others are only recorded on crossbreds (Lutaaya *et al.*, 2002).

CONCLUSIONS

Theoretical features of most linear and non-linear mixed models commonly used in the context of animal breeding and genetic selection have been summarized in this paper. Obviously, all of these mixed models can be set up in different ways depending upon the trait to be evaluated, the population to be analysed, the final goal of the analysis, but also based on the encountered problems. Thus, it is important to properly define the analysis model that will be used thereafter as genetic evaluation model to provide accurate and useful values of the genetic merit of animals, and hence ensure an efficient genetic selection system.

LITERATURE CITED

Ali, T., and L. Schaeffer. 1987. Accounting for covariances among test day milk yields in dairy cows. Can. J. Anim. Sci. 67:637–644.

Arnold, J. W., J. K. Bertrand, and L. L. Benyshek. 1992. Animal model for genetic evaluation of multibreed data. J. Anim. Sci. 70:3322–3332.

Auvray, B., and N. Gengler. 2002. Feasibility of a Walloon test-day model and study of its potential as tool for selection and management. In: Proceedings of the 2002 Interbull meeting. Vol. 29. Interlaaken, Switzerland. p. 123–127.

Bidanel, J.-P. 1992. Comment exploiter la variabilité génétique entre races: du croisement simple à la souche synthétique. INRA Prod. Anim.:249–254.

Bijma, P. 2006. Estimating maternal genetic effects in livestock. J. Anim. Sci. 84:800–806.

Birchmeier, A. N., R. J. C. Cantet, R. L. Fernando, C. A. Morris, F. Holgado, A. Jara, and M. S. Cristal. 2002. Estimation of segregation variance for birth weight in beef cattle. Livest. Prod. Sci. 76:27–35.

Blöttner, S., B. J. Heins, M. Wensch-Dorendorf, L. B. Hansen, and H. H. Swalve. 2011. Brown Swiss × Holstein crossbreds compared with pure Holsteins for calving traits, body weight, backfat thickness, fertility, and body measurements. J. Dairy Sci. 94:1058–1068.

Boettcher, P. 2005. Breeding for improvement of functional traits in dairy cattle. Ital. J. Anim. Sci. 4:7–16.

Brotherstone, S., I. White, K. Meyer, and others. 2000. Genetic modelling of daily milk yield using orthogonal polynomials and parametric curves. Anim. Sci. 70:407–415.

Bryant, J. R., N. López-Villalobos, J. E. Pryce, C. W. Holmes, D. L. Johnson, and D. J. Garrick. 2007. Short Communication: Effect of Environment on the Expression of Breed and Heterosis Effects for Production Traits. J. Dairy Sci. 90:1548–1553.

Cardoso, F. F., and R. J. Tempelman. 2004. Hierarchical Bayes multiple-breed inference with an application to genetic evaluation of a Nelore-Hereford population. J. Anim. Sci. 82:1589–1601.

Clément, V., B. Bibé, É. Verrier, J.-M. Elsen, E. Manfredi, J. Bouix, and É. Hanocq. 2001. Simulation analysis to test the influence of model adequacy and data structure on the estimation of genetic parameters for traits with direct and maternal effects. Genet. Sel. Evol. 33:1.

Croquet, C., P. Mayeres, A. Gillon, S. Vanderick, and N. Gengler. 2006. Inbreeding depression for global and partial economic indexes, production, type, and functional traits. J. Dairy Sci. 89:2257–2267.

Cunningham, E. P., and O. Syrstad. 1987. Crossbreeding Bos indicus and Bos taurus for milk production in the tropics.

Dal Zotto, R., M. Penasa, M. De Marchi, M. Cassandro, N. López-Villalobos, and G. Bittante. 2009. Use of crossbreeding with beef bulls in dairy herds: effect on age, body weight, price, and market value of calves sold at livestock auctions. J. Anim. Sci. 87:3053–3059.

Dal Zotto, R., M. Penasa, M. Povinelli, and G. Bittante. 2007. Effect of crossbreeding on market value of calves from dairy cows. Ital. J. Anim. Sci. 6:102–104.

Dekkers, J. C. M. 2007. Marker-assisted selection for commercial crossbred performance. J. Anim. Sci. 85:2104–2114.

Dodenhoff, J., L. D. Van Vleck, and D. Wilson. 1999. Comparison of models to estimate genetic effects of weaning weight of Angus cattle. J. Anim. Sci. 77:3176–3184.

Dufrasne, M., M. Rustin, V. Jaspart, J. Wavreile, and N. Gengler. 2011. Using test station and on-farm data for the genetic evaluation of Pietrain boars used on Landrace sows for growth performance. J. Anim. Sci. 89:3872–3880.

Eaglen, S. A. E., M. P. Coffey, J. A. Woolliams, and E. Wall. 2012. Evaluating alternate models to estimate genetic parameters of calving traits in United Kingdom Holstein-Friesian dairy cattle. Genet. Sel. Evol. 44:1–13.

Elzo, M. 1986. Inverse of single trait additive genetic covariance matrix with unequal variances across additive genetic groups. J. Dairy Sci. 69:569–574.

Emmerling, R., M. H. Lidauer, and E. A. Mäntysaari. 2002. Multiple Lactation Random Regression Test-day Model for Simmental and Brown Swiss in Germany and Austria. In: Proceedings of the 2002 Interbull meeting. Vol. 29. Interlaaken, Switzerland. p. 111–117. Available from: https://journal.interbull.org/index.php/ib/article/download/743/735

Famula, T., E. Pollak, and L. D. Van Vleck. 1983. Genetic groups in dairy sire evaluation under a selection model. J. Dairy Sci. 66:927–934.

García-Cortés, L. A., and M. Á. Toro. 2006. Multibreed analysis by splitting the breeding values. Genet. Sel. Evol. 38:601–615.

Gengler, N. 1996. Persistency of lactation yields: A review. Interbull Bull. Available from: https://journal.interbull.org/index.php/ib/article/viewFile/254/254

Gengler, N. 2002. Considerations on the path from test-day records to national and international genetic evaluations and its consequences. In: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, Montpellier, France, August, 2002. Session 1. Institut National de la Recherche Agronomique (INRA). p. 1–4.

Gengler, N., A. Tijani, G. Wiggans, and I. Misztal. 1999. Estimation of (co) variance function coefficients for test day yield with a expectation-maximization restricted maximum likelihood algorithm. J. Dairy Sci. 82:1849–e1.

Gianola, D. 1982. Theory and analysis of threshold characters. J. Anim. Sci. 54:1079–1096.

Gianola, D., and J. L. Foulley. 1983. Sire evaluation for ordered categorical data with a threshold model. Genet. Sel. Evol. 15:1.

Gilmour, A., R. Anderson, and A. Rae. 1985. The analysis of binomial data by a generalized linear mixed model. Biometrika 72:593–599.

Gutiérrez, J. P., I. Fernandez, I. Alvarez, L. J. Royo, and F. Goyache. 2006. Sire X contemporary group interactions for birth weight and preweaning growth traits in the Asturiana de los Valles beef cattle breed. Livest. Sci. 99:61–68.

Hansen, L. 2006. Monitoring the worldwide genetic supply for dairy cattle with emphasis on managing crossbreeding and inbreeding. In: Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006. Instituto Prociência. p. 1–1.

Harris, B. L., J. M. Clark, and R. G. Jackson. 1996. Across breed evaluation of dairy cattle. In: Proceedings of the New Zealand Society of Animal Production. Vol. 56. p. 12–15.

Harville, D. A., and R. W. Mee. 1984. A Mixed-Model Procedure for Analyzing Ordered Categorical Data. Biometrics 40:393.

Havenstein, G., P. Ferket, and M. Qureshi. 2003. Carcass composition and yield of 1957 versus 2001 broilers when fed representative 1957 and 2001 broiler diets. Poult. Sci. 82:1509–1518.

Heins, B., L. Hansen, and A. Seykora. 2006a. Calving difficulty and stillbirths of pure Holsteins versus crossbreds of Holstein with Normande, Montbeliarde, and Scandinavian Red. J. Dairy Sci. 89:2805–2810.

Heins, B., L. Hansen, and A. Seykora. 2006b. Fertility and survival of pure Holsteins versus crossbreds of Holstein with Normande, Montbeliarde, and Scandinavian Red. J. Dairy Sci. 89:4944–4951.

Heins, B. J., L. B. Hansen, and A. De Vries. 2012. Survival, lifetime production, and profitability of Normande × Holstein, Montbéliarde × Holstein, and Scandinavian Red × Holstein crossbreds versus pure Holsteins. J. Dairy Sci. 95:1011–1021.

Henderson, C. 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32:69–83.

Henderson, C. R. 1973. SIRE EVALUATION AND GENETIC TRENDS. J. Anim. Sci. 1973. Available from: http://dx.doi.org//1973.1973Symposium10x

Henderson, C. R. 1975. Best Linear Unbiased Estimation and Prediction under a Selection Model. Biometrics 31:423–447.

Hoeschele, I., and P. VanRaden. 1991. Rapid inversion of dominance relationship matrices for noninbred populations by including sire by dam subclass effects. J. Dairy Sci. 74:557–569.

Ibáñẽz-Escriche, N., J. Reixach, N. Lleonart, and J. L. Noguera. 2011. Genetic evaluation combining purebred and crossbred data in a pig breeding scheme. J. Anim. Sci. 89:3881–3889.

Jamrozik, J., L. R. Schaeffer, and J. C. M. Dekkers. 1997. Genetic Evaluation of Dairy Cattle Using Test Day Yields and Random Regression Model. J. Dairy Sci. 80:1217–1226.

Jensen, J. 2001. Genetic Evaluation of Dairy Cattle Using Test-Day Models. J. Dairy Sci. 84:2803–2812.

Jones, L., M. Goddard, and others. 1990. Five years experience with the animal model for dairy evaluations in Australia. In: Proceedings of the 4th World Congress on Genetics applied to Livestock Production, Edinburgh 23-27 July 1990. XIII. Plenary lectures, molecular genetics and mapping, selection, prediction and estimation. p. 382–385.

Kadarmideen, H., R. Thompson, G. Simm, and others. 2000. Linear and threshold model genetic parameters for disease, fertility and milk production in dairy cattle. BSAS Occas. Publ.:83–84.

Koerhuis, A. N. M., and R. Thompson. 1997. Models to estimate maternal effects for juvenile body weight in broiler chickens. Genet. Sel. Evol. 29:225–249.

Konstantinov, K. V., K. T. Beard, M. E. Goddard, and J. H. J. van der Werf. 2009. Genetic evaluation of Australian dairy cattle for somatic cell scores using multi-trait random regression test-day model. J. Anim. Breed. Genet. 126:209–215.

Lande, R. 1981. The minimum number of genes contributing to quantitative variation between and within populations. Genetics 99:541–553.

Legarra, A., J. K. Bertrand, T. Strabel, R. L. Sapp, J. P. Sánchez, and I. Misztal. 2007. Multi-breed genetic evaluation in a Gelbvieh population. J. Anim. Breed. Genet. 124:286–295.

Lo, L., R. Fernando, R. Cantet, and M. Grossman. 1995. Theory for modelling means and covariances in a two-breed population with dominance inheritance. Theor. Appl. Genet. 90:49–62.

Lo, L. L., R. L. Fernando, and M. Grossman. 1993. Covariance between relatives in multibreed populations: additive model. Theor. Appl. Genet. 87:423–430.

Lo, L. L., R. L. Fernando, and M. Grossman. 1997. Genetic evaluation by BLUP in two-breed terminal crossbreeding systems under dominance. J. Anim. Sci. 75:2877–2884.

Lutaaya, E., I. Misztal, J. W. Mabry, T. Short, H. H. Timm, and R. Holzbauer. 2001. Genetic parameter estimates from joint evaluation of purebreds and crossbreds in swine using the crossbred model. J. Anim. Sci. 79:3002–3007.

Lutaaya, E., I. Misztal, J. W. Mabry, T. Short, H. H. Timm, and R. Holzbauer. 2002. Joint evaluation of purebreds and crossbreds in swine. J. Anim. Sci. 80:2263–2266.

Maniatis, N., and G. E. Pollott. 2003. The impact of data structure on genetic (co) variance components of early growth in sheep, estimated using an animal model with maternal effects. J. Anim. Sci. 81:101–108.

Mayeres, P., J. Stoll, J. Bormann, R. Reents, and N. Gengler. 2004. Prediction of Daily Milk, Fat, and Protein Production by a Random Regression Test-Day Model. J. Dairy Sci. 87:1925–1933.

McAllister, A. J. 2002. Is crossbreeding the answer to questions of dairy breed utilization? J. Dairy Sci. 85:2352–2357.

McDowell, R., J. Wilk, and C. Talbott. 1996. Economic viability of crosses of Bos taurus and Bos indicus for dairying in warm climates. J. Dairy Sci. 79:1292–1303.

Miglior, F., B. Muir, and B. Van Doormaal. 2005. Selection indices in Holstein cattle of various countries. J. Dairy Sci. 88:1255–1263.

Miglior, F., B. Van Doormaal, and G. Kistemaker. 2008. Phenotypic analysis of inbreeding depression for traits measured in Canadian dairy cattle breeds. In: 2008 Annual Meeting of the Canadian Society of Animal Science (CSAS). Canada.

Misztal, I., D. Gianola, and J. L. Foulley. 1989. Computing aspects of a non-linear method of sire evaluation for categorical data. J. Dairy Sci. 72:1557–1568.

Muir, B., G. Kistemaker, J. Jamrozik, and F. Canavesi. 2007. Genetic parameters for a multiple-trait multiple-lactation random regression test-day model in Italian Holsteins. J. Dairy Sci. 90:1564–1574.

Munilla Leguizamon, S., and R. J. Cantet. 2010. Equivalence of multibreed animal models and hierarchical Bayes analysis for maternally influenced traits. Genet. Sel. Evol. 42:20.

Newman, S., and A. Reverter. 2000. Practical aspects of crossbred genetic evaluation. Hybrid Breed. Genet. For. Trees Qld. Dep. Prim. Ind. Brisb. Noosa Qld. Aust.:303–315.

Newman, S., A. Reverter, and D. J. Johnston. 2002. Purebred-crossbred performance and genetic evaluation of postweaning growth and carcass traits in Bos indicus$\times$ Bos taurus crosses in Australia. J. Anim. Sci. 80:1801–1808.

Nielsen, H. M., L. G. Christensen, and J. Ødeg\a ard. 2006. A method to define breeding goals for sustainable dairy cattle production. J. Dairy Sci. 89:3615–3625.

Oltenacu, P. A., and D. M. Broom. 2010. The impact of genetic selection for increased milk yield on the welfare of dairy cows. Anim. Welf. 19:39–49.

Penasa, M., M. de Marchi, R. dal Zotto, A. Cecchinato, M. Cassandro, G. Bittante, L. Bailoni, A. dalle Zotte, R. Mantovani, and others. 2009. Influence of the sire on market value of Belgian Blue$\times$ Brown Swiss crossbred calves. In: Italian Journal of Animal Science. Vol. 8. Edizioni Avenue Media. p. 113–115.

Pollak, E. J., and R. L. Quaas. 1998. Multibreed genetic evaluations of beef cattle. In: Proc 6th World Congr Genet Appl Livest Prod. Vol. 23. p. 81–88. Available from: http://www.bifconference.com/bif2005/pdfs/Pollak.pdf

Ptak, E., and L. R. Schaeffer. 1993. Use of test day yields for genetic evaluation of dairy sires and cows. Livest. Prod. Sci. 34:23–34.

Quaas, R. 1976. Computing the diagonal elements and inverse of a large numerator relationship matrix. Biometrics 32:949–953.

Quaas, R. 1988. Additive genetic model with groups and relationships. J. Dairy Sci. 71:1338–1345.

Quaas, R., and E. Pollak. 1981. Modified equations for sire models with groups. J. Dairy Sci. 64:1868–1872.

Robinson, D. L. 1996. Models which might explain negative correlations between direct and maternal genetic effects. Livest. Prod. Sci. 45:111–122.

Robinson, G. 1986. Group effects and computing strategies for models for estimating breeding values. J. Dairy Sci. 69:3106–3111.

de Roos, A. P. W., A. G. F. Harbers, and G. de Jong. 2004. Random Herd Curves in a Test-Day Model for Milk, Fat, and Protein Production of Dairy Cattle in The Netherlands. J. Dairy Sci. 87:2693–2701.

Sánchez, J. P., I. Misztal, I. Aguilar, and J. K. Bertrand. 2008. Genetic evaluation of growth in a multibreed beef cattle population using random regression-linear spline models. J. Anim. Sci. 86:267–277.

Schaeffer, L. R. 2004. Application of random regression models in animal breeding. Livest. Prod. Sci. 86:35–45.

Schaeffer, L. R., and J. C. M. Dekkers. 1994. Random regressions in animal models for test-day production in dairy cattle. In: Proceedings of Fifth World Congress of Genetics Applied to Livestock Production. Vol. 18. p. 443–446.

Schaeffer, L. R., J. Jamrozik, G. J. Kistemaker, and B. J. Van Doormaal. 2000. Experience with a test-day model. J. Dairy Sci. 83:1135–1144.

Sørensen, A. C., M. K. Sørensen, and P. Berg. 2005. Inbreeding in Danish dairy cattle breeds. J. Dairy Sci. 88:1865–1872.

Sørensen, M. K., E. Norberg, J. Pedersen, and L. G. Christensen. 2008. Invited Review: Crossbreeding in Dairy Cattle: A Danish Perspective. J. Dairy Sci. 91:4116–4128.

Stanton, T., L. Jones, R. Everett, and S. Kachman. 1992. Estimating milk, fat, and protein lactation curves with a test day model. J. Dairy Sci. 75:1691–1700.

Sullivan, P. G., J. W. Wilton, S. P. Miller, and L. R. Banks. 1999. Genetic trends and breed overlap derived from multiple-breed genetic evaluations of beef cattle for growth traits. J. Anim. Sci. 77:2019–2027.

Swalve, H. H. 1995. The Effect of Test Day Models on the Estimation of Genetic Parameters and Breeding Values for Dairy Yield Traits. J. Dairy Sci. 78:929–938.

Swalve, H. H. 1998. Use of test day records for genetic evaluation. In: Proc 6th World Congr Genet Appl Livest Prod. Vol. 23. Armidale, Australia. p. 295–302.

Swalve, H. H. 2000. Theoretical Basis and Computational Methods for Different Test-Day Genetic Evaluation Methods. J. Dairy Sci. 83:1115–1124.

Swan, A. A., and B. P. Kinghorn. 1992. Evaluation and exploitation of crossbreeding in dairy cattle. J. Dairy Sci. 75:624–639.

Thompson, R. 1979. Sire Evaluation. Biometrics 35:339.

Van Der Werf, J. H. J., and W. De Boer. 1989. Influence of Nonadditive Effects on Estimation of Genetic Parameters in Dairy Cattle. J. Dairy Sci. 72:2606–2614.

Van Tassell, C., R. Quaas, and R. Everett. 1992. Parameter estimates for 305-day ME records and 305-day test-day residual records. J Dairy Sci 75:251.

Van Vleck, L. D. 1997. Calculation of breed direct and maternal genetic fractions and breed specific direct and maternal heterozygosity for crossbreeding data. Braz. J. Genet. 20. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-84551997000400013&lng=en&nrm=iso&tlng=en

VanRaden, P., and I. Hoeschele. 1991. Rapid inversion of additive by additive relationship matrices by including sire-dam combination effects. J. Dairy Sci. 74:570–579.

VanRaden, P. M., and A. H. Sanders. 2003. Economic merit of crossbred and purebred US dairy cattle. J. Dairy Sci. 86:1036–1044.

VanRaden, P. M., M. E. Tooker, J. B. Cole, G. R. Wiggans, and J. H. Megonigal. 2007. Genetic Evaluations for Mixed-Breed Populations. J. Dairy Sci. 90:2434–2441.

Walsh, S. W., E. J. Williams, and A. C. O. Evans. 2011. A review of the causes of poor fertility in high milk producing dairy cows. Anim. Reprod. Sci. 123:127–138.

Wei, M., J. H. J. Werf, and E. W. Brascamp. 1991. Relationship between purebred and crossbred parameters: II Genetic correlation between purebred and crossbred performance under the model with two loci. J. Anim. Breed. Genet. 108:262–269.

Wei, M., and J. H. J. van der Werf. 1994. Maximizing genetic response in crossbreds using both purebred and crossbred information. Anim. Prod. 59:401–413.

Wei, M., and J. H. van der Werf. 1995. Genetic correlation and heritabilities for purebred and crossbred performance in poultry egg production traits. J. Anim. Sci. 73:2220–2226.

Weigel, K. A., and K. A. Barlass. 2003. Results of a producer survey regarding crossbreeding on US dairy farms. J. Dairy Sci. 86:4148–4154.

van der Werf, J. 2002. Mixed models for genetic analysis. In: The attached file of ASReml User Guide. University of New England Armidale, Australia. Available from: http://vsn.klever.co.uk/products/asreml/user/geneticanalysis.pdf

Westell, R., R. Quaas, and L. D. Van Vleck. 1988. Genetic groups in an animal model. J. Dairy Sci. 71:1310–1318.

Willham, R. L. 1963. The Covariance between Relatives for Characters Composed of Components Contributed by Related Individuals. Biometrics 19:18.

Willham, R. L. 1972. The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. J. Anim. Sci. 35:1288–1293.

Wilmink, J. 1987. Studies on test-day and lactation milk, fat and protein yield of dairy cows [Ph.D. Thesis]. Landbouwuniversiteit te Wageningen, Netherlands.

Wolfová, M., J. Wolf, J. Kvapilík, and J. Kica. 2007. Selection for profit in cattle: II. Economic weights for dairy and beef sires in crossbreeding systems. J. Dairy Sci. 90:2456–2467.

Wright, S. 1934. An analysis of variability in number of digits in an inbred strain of guinea pigs. Genetics 19:506–536.

Yang, N., and R.-S. Jiang. 2005. Recent advances in breeding for quality chickens. Worlds Poult. Sci. J. 61:373–381.

Zumbach, B., I. Misztal, S. Tsuruta, J. Holl, W. Herring, and T. Long. 2006. Genetic correlations between two strains of Durocs and crossbreds from differing production environments for slaughter traits. J. Anim. Sci. 85:901–908.

CHAPTER III. DEVELOPMENT OF GENETIC EVALUATION SYSTEMS FOR MATERNALLY INFLUENCED TRAITS

CHAPTER III. DEVELOPMENT OF GENETIC EVALUATION SYSTEMS FOR MATERNALLY INFLUENCED TRAITS

OUTLINE

As discussed in Chapter II, when traits are influenced not only by the genotype and/or the environment of the individual itself, but also by either the genotype and/or environment from its mother, these traits are denoted as being maternally influenced traits, i.e. traits influenced by maternal effects. All of these effects need to be considered in models used for the genetic evaluation of those maternally influenced traits. It was stated in Chapter II that the maternal animal model enables the estimation of direct additive genetic as well as maternal additive genetic and environmental effects affecting those traits. As first step in the research strategy of this thesis, this chapter focused on the development of genetic evaluation systems for maternally influenced traits using appropriate maternal animal models to provide accurate and useful breeding values for both additive effects. To this end, two studies were undertaken on two maternally influenced traits: the first was conducted using data of calving ease scores from Walloon Holstein dairy cattle (Chapters III.1 and III.2), and the second was conducted using data of lamb survival to weaning from New Zealand sheep (Chapter III.3).

CHAPTER III.1 GENETIC PARAMETERS FOR DIRECT AND MATERNAL CALVING EASE IN WALLOON DAIRY CATTLE BASED ON LINEAR AND THRESHOLD MODELS

Based on: Vanderick, S., T. Troch, A. Gillon, G. Glorieux, and N. Gengler. 2014. Genetic parameters for direct and maternal calving ease in Walloon dairy cattle based on linear and threshold models. J. Anim. Breed. Genet. 131:513–521.

CHAPTER III.1 GENETIC PARAMETERS FOR DIRECT AND MATERNAL CALVING EASE IN WALLOON DAIRY CATTLE BASED ON LINEAR AND THRESHOLD MODELS

OUTLINE

Calving ease is an important functional trait in dairy cattle because it is related to the production, fertility and general health during the subsequent lactation, and hence can affect negatively economic profitability if complications arise at calving time. Until recently, there was no genetic evaluation of this trait in Wallonia. Therefore; there was an obvious consensus that a genetic evaluation system for calving ease in Walloon dairy cattle was needed. As calving ease is a maternally influenced trait two individuals need to be considered, the cow and her calf, *i.e.* the maternal and the direct effects of calving ease, respectively. Thus, a maternal animal model as described in Chapter II was required. Moreover, according to the categorical nature of this trait, theoretically, a threshold model is more appropriate than linear model, though in practice it seems there is little to be gained by its use. Therefore, the objective of the present chapter was to assess the use of linear and threshold maternal animal models to analyse calving ease scores from the Holstein dairy cattle in the Walloon Region of Belgium. First, (co) variance components and derived genetic parameters for direct and maternal additive effects were estimated by using linear and threshold maternal animal models. Then, the relative merits of models were compared in terms of predictive ability.

ABSTRACT

Calving ease scores from Holstein dairy cattle in the Walloon Region of Belgium were analysed using univariate linear and threshold animal models. Variance components and derived genetic parameters were estimated from a data set including 33 155 calving records. Included in the models were season, herd and sex of calf × age of dam classes × group of calvings interaction as fixed effects, herd × year of calving, maternal permanent environment and animal direct and maternal additive genetic as random effects. Models were fitted with the genetic correlation between direct and maternal additive genetic effects either estimated or constrained to zero. Direct heritability (h^2) for calving ease was about 8% with linear models and about 12% with threshold models. Maternal heritabilities were about 2% and 4%, respectively. Genetic correlation between direct and maternal additive effects was found to be not significantly different from zero. Models were compared in terms of goodness of fit and predictive ability. Criteria of comparison such as mean squared error, correlation between observed and predicted calving ease scores as well as between estimated breeding values (EBV) were estimated from 85 118 calving records. The results provided few differences between linear and threshold models even though correlations between EBV from subsets of data for sires with progeny from linear model were 17% and 23 % greater for direct and maternal genetic effects, respectively, than from threshold model. For the purpose of genetic evaluation for calving ease in Walloon Holstein dairy cattle, the linear animal model without covariance between direct and maternal additive effects was found to be the best choice.

Keywords: dystocia, animal model, Holstein, heritability

INTRODUCTION

All dairy cows must give birth in order to begin producing milk. In most cases, calving proceeds normally but problems may happen before or during the calving and cause various problems. The major problem is dystocia which may be defined as calving difficulty resulting from prolonged spontaneous calving or prolonged or severe assisted extraction. Mee (2008) provides a good review of the different types of dystocia and their associated risk factors in dairy cattle. Calving complications impact production, fertility, and cow and calf morbidity and mortality and thus can negatively affect economic profitability in dairy herds (Dekkers, 1994; Dematawena and Berger, 1997; López de Maturana *et al.*, 2007b; Eaglen *et al.*, 2011). Calving-related infections affect also indirectly human health as they require increased use of antibiotics, leading potentially to microbial resistance. Besides, animal welfare is compromised by these calving complications and so consumer acceptability of dairy production systems (Mee, 2008).

Calving ease measures the presence or absence of dystocia and its intensity. This trait is generally scored on a categorical scale by the breeder which makes it more sensitive to subjectivity (Dekkers, 1994). Furthermore, this trait is affected by two additive genetic components, the calf's contribution (direct effect; *e.g.* arising from size, birth weight, hormonal balance…) and the dam's contribution (maternal effect; *e.g.* arising from pelvic opening, uterine influence of the dam on her calf's birth weight…). The direct additive effect is expressed only once, when the calf is born, whereas the maternal additive effect is expressed several times, each time a cow calves.

From a theoretical point of view, threshold models are preferred over linear models as a method for genetic analysis of such categorical traits displaying a discrete probability distribution (Gianola, 1982) and this was confirmed with simulated data by Hoeschele (1988). However, several studies in sheep, beef and dairy cattle using field data found no clear advantage of threshold over linear models (Weller and Gianola, 1989; Olesen *et al.*, 1994; Matos *et al.*, 1997; Varona *et al.*, 1999; Ramirez-Valverde *et al.*, 2001; Phocas and Laloë, 2003). Some of these studies reported greater computational requirements with threshold than with linear models. This might explain why most of the routine genetic evaluations of categorical calving traits are based on linear models (Interbull, 2013) although such data violate the assumption of normality. Calving traits are evaluated with a threshold approach only in France, Italy and the USA (Ducrocq, 2000; Canavesi *et al*., 2003; Wiggans *et al.*, 2003).

Models used for routine genetic evaluation of calving ease range from sire (-maternal grandsire) models to animal models in univariate or multiple-trait form that either allow a covariance between direct and maternal genetic effects or fix this covariance to zero (Interbull, 2013). Many threshold models are implemented as models with sire-maternal grandsire effects to avoid convergence problems and biased estimation of genetic parameters due to the well-known extreme category problem, particularly in the presence of numerous fixed effect classes (Luo *et al*., 2001). However, because some cows with calving records that also have their own direct calving records as a calf, an animal model seems more appropriate to include information on the cows themselves and so generates directly EBV for direct and maternal effects for bulls and cows.

The purpose of this research was to compare linear and threshold animal models for the prediction of EBV for calving ease and to estimate the genetic parameters for direct and maternal additive effects for calving ease in the Walloon Holstein dairy cattle. Models were compared on the basis of their predictive abilities to determine the most suitable model for current Walloon data.

MATERIALS AND METHODS

DATA

In the Walloon Region of Belgium, calving ease is scored by dairy breeders on a voluntary basis and collected by the Walloon Breeding Association (AWE). Calving ease scores range from 1 to 4 (1. caesarean and embryotomy, 2. hard pull, 3. easy pull, and 4. normal). The original data set comprised 138 144 calving records and presented a typical distribution of calving ease; most of the records fell into category 4 (69%) and few records into category 1 (approximately 1%). From this original data set, two data sets were created, one for the (co)variance components estimation (data set I) and one for the validation/comparison of models (data set II). Records from Holstein calves born between 2000 and 2012 were used for this research and data editing was almost identical for both data sets.

Data were edited to remove all suspect records, which included records with out-of-range values for calving ease or missing information related to the factors in the statistical model, including animal identification, birth date, herd identification, calving date, parity number, calving scores and sex of calf. Only records on single born calves were used. Records were limited to first five parities. Calving age of dams was restricted to be between 21 and 48 months for primiparous (1st parity) cows and between 31 and 142 months for multiparous (2nd to 5th parities) cows. Percentage of records for dams outside these ages was relatively small (<0.2%). Data quality depends highly on dairy breeders' own judgement to assign scores for calving ease. Therefore, only herds with a standard deviation for scores >0.05 were kept to avoid herds where breeders put all scores in the same category. In addition to all the general edits, some specific edits were applied to each data set.

For the data set I, all calves had to have sire and dam identified and every dam had to have a calving record in first parity. Herds displaying less than four first calvings on average per year were deleted. In each herd, only data from continuous calvings per dam were kept (*e.g.* if a dam displayed records from its first, second and fourth calvings only records from first and second calvings were kept). A final edit required on average more than one calving per dam per herd. The objective was to create a reliable data set without unnecessarily reducing the available data. The final data set I included records from 33 155 calves born in 492 Walloon herds from 2215 sires, 25 240 dams and 2031 maternal grandsires. The total number of animals including ancestors without records was 120 374.

For the data set II, all calves had to have only dam identified and herds had to display at least, on average, four calvings per year calculated from the first two parities. The final data set II included records from 85 118 calves originating from 862 Walloon herds, from 3148 sires, 62 265 dams and 3352 maternal grandsires. The total number of animals in the pedigree was 233 882.

For both data sets, calving ages of dam were divided into eleven classes: 21–24, 25–26, 27–28, 29–30, 31–35, 36–38, 39–48, 49–56, 57–65, 66–81 and more than 81 months at calving. Calving seasons were divided into four seasons: winter season from January to March, spring season from April to June, summer season from July to September and autumn season from October to December.

MODELS OF ANALYSIS

All the fitted models included the three following fixed effects: season effects, herd effects and combined effects of sex of calf by age of dam classes by group of parities (two groups: first parity and the 2nd to the 5th parity).

Univariate Linear Animal Model

Calving ease was modelled as a continuous trait:

$$\mathbf{y}_{CE} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_h\mathbf{h} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_m\mathbf{m} + \mathbf{Z}_p\mathbf{p} + \mathbf{e} \qquad (38)$$

where $\mathbf{y}_{CE}$ is a vector of observed calving ease scores, $\boldsymbol{\beta}$ is a vector of fixed effects; $\mathbf{h}$ is a vector of random herd × year of calving effects which were included to account for the variability in the frequency of dystocia among herds and years within herds, $\mathbf{a}$ is a vector of random direct additive genetic effects; $\mathbf{m}$ is a vector of random maternal additive genetic effects, $\mathbf{p}$ is a vector of random permanent maternal environmental effects; $\mathbf{X}$, $\mathbf{Z}_h$, $\mathbf{Z}_a$, $\mathbf{Z}_m$

and $\mathbf{Z}_p$ are incidences matrices linking observations with respective effects; and $\mathbf{e}$ is a vector of residual effects. There might be some statistical problems and convergence issues with the estimation of fixed herd × year of calving with a threshold model when dealing with herd × year of small size or with some scores not registered (*i.e.* the extreme category problem). A random herd × year effect was fitted to avoid this problem (Misztal *et al*., 1989). Even if this issue is less a problem in linear models, the same random herd × year effect was kept.

Model indicated as Model L1 was fitted with an estimated genetic correlation between direct and maternal additive genetic effects. Model indicated as Model L2 was fitted with a genetic correlation between genetic effects constrained to zero.

Univariate Threshold Animal Model

The same fixed and random effects as in the linear model were considered, but this model assumed the existence of an underlying unobservable normal variable - that is, a liability (L) - modelling the response of calving ease with the following distribution:

$$\begin{aligned} f(y_{CE}|L) = \prod_{i=1}^{n} f(y_{CE}|L_i) = \prod_{i=1}^{n} \big[& I(L_i < t_1) I(y_{CE\,i} = 1) \\ & + I(t_1 < L_i < t_2) I(y_{CE\,i} = 2) \\ & + I(t_2 < L_i < t_3) I(y_{CE\,i} = 3) \\ & + I(L_i > t_3) I(y_{CE\,i} = 4) \big] \end{aligned} \qquad (39)$$

where y_{CE} are the observed calving ease scores, t_1, t_2 and t_3 are thresholds that categorize the four categories of response and I is an indicator function that takes value 1 if the condition specified is true and 0 otherwise. A response in a given category is observed, if the actual value of liability falls between the thresholds defining the appropriate category.

Just as Wang *et al.* (1997), thresholds t_1 and t_2 were assumed to be known and t_3 was assumed to be unknown in order to simplify the sampling scheme rather than the one defined by setting the residual variance of the categorical trait to one. Therefore, the values of t_1 and t_2 were based on the observed frequencies of calving ease scores in the considered categories, and residual variance was assumed to be unknown.

Model indicated as Model T1 was fitted with an estimated genetic correlation between direct and maternal additive genetic effects. Model indicated as Model T2 was fitted with a genetic correlation between genetic effects constrained to zero.

Variance components were estimated based on the data set I, for the four models by a Bayesian approach using the Gibbs sampling algorithm with flat priors for (co)variances. Gibbs sampling was used to obtain the marginal posterior distribution for variance components of each random effect from the model from 400 000 samples, after discarding 50 000 samples as the burn-in period. The stationary stage was confirmed by graphical inspection of plots of sampled values *vs* iterations. Every fifth sample was retained to compute mean and standard deviation of the marginal posterior distribution. The estimation and the post-Gibbs analysis were performed using programs kindly provided by Ignacy Misztal (Misztal *et al*., 2002).

COMPARISON OF MODELS

Models were compared on their goodness of fit but also on their ability to predict ‘future data’. For this purpose, the entire data set II was split into two parts. One-half of the calving ease records was randomly set to missing in the first data subset and the remaining one-half was set to missing in the second data subset. So, each calving ease record was only present in one of the two subsets. For these two subsets, direct and maternal EBV and expectations of calving ease score were computed with a BLUP approach for linear and threshold models. This strategy was repeated five times in order to get ten subsets*, that is,* five paired subsets.

Within each model and for each of the ten data subsets, mean square errors (MSE) were computed between expectations from the predictive distribution and the observed calving ease records which had been randomly set to missing.

The MSE was defined for linear model as:

$$\mathrm{MSE} = \frac{1}{\mathrm{n}} \sum_{i=1}^{n} \left(y_{CEi} - \hat{y}_{CEi} \right)^2 \qquad (40)$$

where y_{CE} and $\hat{y}_{CE}$ correspond to the observed and predicted calving ease scores, respectively; n is the number of data points in a data subset.

With the threshold model, MSE was computed, based on López de Maturana *et al.*, (2009) as:

$$\mathrm{MSE} = \frac{1}{n} \left(\sum_{i=1}^{n} y_{CEi} - \sum_{C=1}^{ncat} C \times P_{Ci} \right)^2 \qquad (41)$$

where the probability (P_{Ci}) that observation *i* falls in category *C* was computed as:

$$P_{Ci} = \Phi\left(\frac{t_C - L_{CEi}}{\hat{\sigma}_e} \right) - \Phi\left(\frac{t_{C-1} - L_{CEi}}{\hat{\sigma}_e} \right) \qquad (42)$$

with $\Phi(.)$ is the cumulative distribution function of a normal variable evaluated at $(.)$; t_C is the inferred value of the appropriate threshold, and L_{CEi} is the posterior mean of the liability to calving ease for an individual *i*.

Within each model and for each of the ten data subsets, Pearson's correlation between observed and predicted scores was calculated as:

$$r(y_{CE}, \hat{y}_{CE}) = \frac{\mathrm{cov}(y_{CE}, \hat{y}_{CE})}{\sigma_{y_{CE}} \sigma_{\hat{y}_{CE}}} \qquad (43)$$

where $\mathrm{cov}(y_{CE}, \hat{y}_{CE})$ is the estimate of covariance between the observed and predicted calving ease scores and $\sigma_{y_{CE}}$ and $\sigma_{\hat{y}_{CE}}$ are the estimates of standard deviations of observed and predicted calving ease scores, respectively.

Three groups of sires with progeny were created depending on their accuracy as follows: low: sires $>0 \leq 50$ progeny, medium: sires $>50 \leq 100$ progeny and high: sires >100 progeny. For each of these three groups of sires, correlations between sire EBV were calculated for each of the five paired data subsets within each model in order to assess model prediction performance. A higher correlation estimate implied a better stability of the model to predict EBV for animals whose records were randomly set to missing.

Finally, Spearman's rank correlations were computed between sire EBV estimated from linear model and from threshold model for sires with progeny.

RESULTS AND DISCUSSION

The descriptive statistics of the data set I and data set II are displayed in Table 1. Disproportionate sex ratios were observed in records, there were fewer male calves than female calves in both data sets. Further investigation suggested there may be a bias in recording of the sex of the calf, as some breeders prefer to record female calves than male calves due to the difference in value between a male and a female calf in dairy cattle. This recording bias in the number of female and male calves can lead to an underreporting of difficulty to calve (score < 4) since the calving of male calves is known to be more difficult (Mee, 2008).

Table 1. Summary of edited data set used to estimate (co)variance components (data set I) and edited data set used for validation of models (data set II)

	Data set I		Data set II	
Item	No. of observations	Percentage	No. of observations	Percentage
Final data file	33 155	-	85 118	-
Female calves	26 177	78.9	66 511	78.1
Male calves	6978	21.1	18 511	21.9
Final pedigree file	120 374	-	233 882	-
Herds	492	-	862	-
Sires with progeny records	2215	-	3148	-
> 0 ≤ 50 progeny	2067	-	2785	-
> 50 ≤100 progeny	90	-	180	-
> 100 progeny	58	-	183	-
Dams	25 240	-	62 265	-
Maternal grandsires	2031	-	3352	-
Calving ease				
1. Caesarean and embryotomy	443	1.3	781	0.9
2. Hard pull	2179	6.6	4006	4.7
3. Easy pull	10 114	30.5	23 461	27.6
4. Normal	20 419	61.6	56 870	66.8

(CO)VARIANCE COMPONENTS AND DERIVED GENETIC PARAMETERS

Results for the (co)variance components and derived genetic parameters are reported in Table 2. Parameters generally were significantly different from zero because posterior means were more than two posterior standard deviations from zero, except for genetic correlation between direct and maternal additive genetic effects for Model L1 and Model T1 (*i.e.* models fitted with an estimated genetic correlation).

Table 2. Posterior mean (PM) and posterior standard deviation (PSD) of (co)variance components and related genetic parameters estimated with each of four models

	Model L1[a]		Model L2[a]		Model T1[a]		Model T2[a]	
Parameter[b]	PM	PSD	PM	PSD	PM	PSD	PM	PSD
σ_h^2	0.042	0.002	0.042	0.002	0.146	.012	0.146	0.012
σ_a^2	0.027	0.004	0.028	0.004	0.085	0.015	0.082	0.013
σ_m^2	0.008	0.003	0.009	0.002	0.027	0.008	0.024	0.008
σ_p^2	0.018	0.004	0.017	0.005	0.035	0.012	0.034	0.010
σ_e^2	0.269	0.005	0.269	0.005	0.411	0.049	0.413	0.048
$r_u(a,m)$	0.088	0.194	-	-	-0.071	0.190	-	-
h_a^2	0.074	0.012	0.078	0.012	0.121	0.024	0.117	0.020
h_m^2	0.023	0.007	0.024	0.007	0.039	0.012	0.034	0.011
C_h	12%		12%		21%		21%	
C_p	5%		5%		5%		5%	
C_e	74%		74%		59%		59%	

[a] Model L1 is the linear animal model with estimated covariance, Model L2 is the linear animal model with covariance constrained to zero, Model T1 is the threshold animal model with estimated covariance and Model T2 is the threshold animal model with covariance constrained to zero

[b] The terms σ_h^2 is the herd × year of calving variance, σ_a^2 is the direct additive genetic variance, σ_m^2 is the maternal additive genetic variance, σ_p^2 is the permanent maternal environmental variance, σ_e^2 is the residual variance, $r_u(a,m)$ is the genetic correlation between direct and maternal effects, h_a^2 and h_m^2 are the direct and the maternal heritabilities, respectively. C_h, C_p and C_e are the herd × year of calving fraction, permanent maternal environmental fraction and residual fraction in the phenotypic variance, respectively.

The additive genetic variance due to direct effects was greater than that due to maternal effects for all models. On average, direct heritabilities were approximately three to four times as large as maternal heritabilities. All heritabilities estimated with all models were within the range of previously published estimates of this trait in dairy cattle, which ranged from 0.03 to 0.17 for direct h^2 and from 0.02 to 0.12 for maternal h^2 (Weller and Gianola, 1989; Steinbock *et al*., 2003; Wiggans *et al.*, 2003; López de Maturana *et al.*, 2007a; Eaglen *et al.*, 2012). These estimates are not directly comparable because of different models (animal *vs* sire and maternal grandsire, linear *vs* threshold, univariate *vs* bivariate) that were used. However, most previous estimates tended to show that direct h^2 was greater than maternal h^2.

Effects of herd × year of calving represent differences among herds and years of calving which can be partly due to differences in subjective scoring of calving ease. Fitting herd × year effects as random allows more effective use of the data when applying the threshold model. The herd × year of calving effects represented 12 and 21% of the phenotypic variance for the linear and threshold models, respectively, which was the largest contributor to the phenotypic variance after the residual effects (74 and 59%).

The maternal permanent environment effects represented 5% of the phenotypic variance in each model and were greater than the genetic maternal effects. Preliminary analyses based on the current data showed that maternal genetic variances tended to be overestimated by models that ignored permanent environmental effects. Estimates of variance components and derived genetic parameters were similar within model type (linear *vs* threshold). A positive genetic correlation was found with Model L1 and a negative one with Model T1 but in both cases the genetic correlation was not significant. Therefore, it seemed more appropriate to consider no genetic correlation between direct and maternal additive genetic effects in the subsequent stage of this study.

The analysis of calving ease with linear models yielded variance estimates that were consistently smaller than those obtained with threshold models. Particularly, variance of herd × year of calving effects showed a marked decrease from threshold models to linear models. Threshold model h^2 estimates were greater than linear model h^2 estimates (0.117 *vs* 0.078 and 0.034 *vs* 0.024 for direct and maternal heritabilities, respectively) but these heritabilities cannot be directly compared since they were estimated on different scales; on a visible probability scale and on an underlying normal scale for linear and threshold models, respectively. Furthermore, h^2 estimates are frequency dependent when a linear model is used to fit categorical traits. Dempster and Lerner (1950) proposed transformations to make heritabilities comparable. As reported by several studies, higher heritabilities are usually expected with threshold models than linear models (Luo *et al*., 1999; Phocas and Laloë, 2003).

The best fit of the model, measured by the percentage of residual variance in the phenotypic variance, was achieved for threshold models, approximately 59% against 74% with linear models.

COMPARISON OF MODELS

The MSE for Model L2 and Model T2 used to predict the calving ease records set to missing in the ten data subsets is provided in Table 3. Models with the smaller MSE had better predictive ability. In general, MSE was similar for both models with only very small differences. Based on the average MSE, the threshold model did not perform better than linear model (0.294 *vs* 0.293). These results were consistent with those obtained by Varona *et al.* (1999) who also used differences in MSE as a criterion for comparison of models. They found small differences between univariate linear and threshold models based on field and simulated data in beef cattle.

Table 3 also displays Pearson’s correlation estimates between observed and predicted calving ease scores by Models L2 and T2 for the ten subsets. Similar to MSE, differences in correlation between models were very small. For all subsets, the threshold model performed slightly better than the linear model (0.502 *vs* 0.497). These results were expected since the threshold model is considered as being strategy better model to fit such categorical traits.

Table 3. Mean squared error (MSE) and Pearson's correlation estimates between observed and predicted calving ease scores for Model L2 and Model T2 for the ten replicates

	MSE		Correlation	
Subset	Model L2[a]	Model T2[a]	Model L2[a]	Model T2[a]
1	0.292	0.292	0.497	0.502
2	0.296	0.297	0.493	0.497
3	0.295	0.294	0.497	0.503
4	0.292	0.293	0.495	0.499
5	0.291	0.292	0.501	0.507
6	0.294	0.295	0.494	0.498
7	0.293	0.293	0.494	0.500
8	0.292	0.293	0.501	0.506
9	0.292	0.292	0.494	0.500
10	0.294	0.294	0.499	0.505
Average	0.293	0.294	0.497	0.502

[a] Model L2 is the linear animal model with covariance constrained to zero and Model T2 is the threshold animal model with covariance constrained to zero

Table 4 contains the average correlation estimates between the five paired data subsets for genetic direct and maternal calving ease EBV from Model L2 and Model T2 considering sires with 50 or fewer progeny (low-accuracy sires), sires with between 51 and 100 progeny (medium-accuracy sires), and sires with more than 100 progeny (high-accuracy sires). The differences between linear and threshold models decreased as the number of progeny records available for sires increased, especially for differences between sire EBV for maternal effects. So, if the number of calving records per sire is limited, differences in the ranking of sires might occur using the linear *vs* threshold model. For all groups of sires, greater correlations were observed with Model L2 than with Model T2 for direct and maternal genetic effects. On average, correlations from the linear model were 17 and 23% higher than from the threshold model for direct and maternal EBV, respectively. Thus, the linear model appeared to have a higher stability for predicting EBV of animals whose records were randomly set to missing. These results were not in line with those obtained in beef cattle by Ramirez-Valverde *et al.*

(2001) who found a better stability with a threshold approach. As expected, lower accuracy was observed for maternal effects.

Table 4. Average, standard deviation (SD), minimum and maximum of correlation estimates between split data sets[a] for calving ease breeding values of sires with progeny (N=3148) from Model L2 and Model T2

	Model L2[c]							
	Direct genetic effect				Maternal genetic effect			
Category[b]	Mean	SD	Min	Max	Mean	SD	Min	Max
Sires > 0 ≤ 50	0.635	0.028	0.604	0.663	0.465	0.057	0.369	0.506
Sires > 50 ≤100	0.647	0.047	0.605	0.698	0.369	0.037	0.329	0.407
Sires > 100	0.689	0.030	0.644	0.721	0.394	0.073	0.290	0.460
	Model T2[c]							
Sires > 0 ≤ 50	0.507	0.030	0.475	0.536	0.360	0.047	0.279	0.392
Sires > 50 ≤100	0.549	0.057	0.478	0.610	0.282	0.032	0.238	0.316
Sires > 100	0.599	0.038	0.538	0.634	0.334	0.069	0.226	0.406

[a] In five paired data subsets.

[b] Sires > 0 ≤ 50: sires with 50 or fewer progeny records in data file, sires > 50 ≤100: sires with 51-100 progeny records in data file, sires > 100: sires with more than 100 progeny records in data file.

[c] Model L2 is the linear animal model with covariance constrained to zero and Model T2 is the threshold animal model with covariance constrained to zero

The most likely reasons for the linear model showing consistently better results could be due to the fact that in the threshold model additional parameters (thresholds) needs to be estimated leading potentially to lower estimation accuracies, especially for animal models. The threshold model fitted slightly better, explained more variance; however, EBV were less stable between paired subsets especially for maternal additive genetic effects.

Fitting herd × year effect as random can lead to biased estimates of EBV (Visscher and Goddard, 1993). Phocas and Laloë (2003) stated that when a non-random association exists between sires and contemporary groups, the correlation between true and predicted EBV can be affected. However, it is unsure to what degree this non-random association has to exist to create this behaviour.

Spearman's rank correlations between sire EBV from Model L2 and Model T2 were 0.972 and 0.971 for direct and maternal calving ease EBV, respectively, indicating that the ranking of sires were nearly identical between the linear and the threshold models. This was in agreement with results from similar comparisons involving categorical traits in cattle (Weller *et al.*, 1988; Clutter *et al.*, 1989; Ramirez-Valverde *et al.*, 2001) and in sheep (Olesen *et al*., 1994; Matos *et al.*, 1997).

CONCLUSIONS

(Co)variance components and derived genetic parameters for calving ease were estimated with univariate linear and threshold animal models. The direct-maternal genetic correlation was positive for the linear model and negative for the threshold model, but neither was significantly different from zero. The h^2 estimates were consistent with those found in other studies on calving ease in dairy cattle. The threshold models showed a better goodness of fit than linear models. However, in terms of predictive ability, no clear advantage of the threshold models over the linear models was found with our data. Accordingly, it would be preferable and more technically feasible to use a linear model to perform genetic evaluation of calving ease. Thus, the linear animal model without covariance between direct and maternal additive genetic effects (*i.e.* Model L2) would be the model of choice to implement the routine genetic evaluation of calving ease for the Walloon dairy cattle.

ACKNOWLEDGMENTS

The Ministry of Agriculture of the Walloon Region of Belgium (Service Public de Wallonie, Direction générale opérationnelle "Agriculture, Ressources naturelles et Environnement) is acknowledged for its financial support. The authors gratefully acknowledge the technical support provided by the Walloon Breeding Association (Ciney, Belgium). Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI) funded by the National Fund for Scientific Research (F.R.S.-FNRS, Brussels, Belgium) under Grant No. 2.5020.11. The authors are grateful to the anonymous reviewers for their helpful suggestions and comments on the manuscript.

LITERATURE CITED

Canavesi, F., S. Biffani, and A.B. Samoré. 2003. Revising the genetic evaluation for calving ease in the Italian Holstein Friesian. Interbull Bull. 30:82.

Clutter, A.C., P.J. Berger, and J.M. Mattison. 1989. Threshold model analysis of dystocia in dairy cattle when progeny information is limited. J. Dairy Sci. 72:3264–3272.

Dekkers, J.C.M. 1994. Optimal breeding strategies for calving ease. J. Dairy Sci. 77:3441–3453.

Dematawena, C.M.B., and P.J. Berger. 1997. Effect of dystocia on yield, fertility, and cow losses and an economic evaluation of dystocia scores for Holsteins. J. Dairy Sci. 80:754–761.

Dempster, E.R., and I.M. Lerner. 1950. Heritability of threshold characters. Genetics. 35:212.

Ducrocq, V. 2000. Calving ease evaluation of French dairy bulls with a heteroskedastic threshold model with direct and maternal effects. Interbull Bull. 25:123.

Eaglen, S.A.E., M.P. Coffey, J.A. Woolliams, R. Mrode, and E. Wall. 2011. Phenotypic effects of calving ease on the subsequent fertility and milk production of dam and calf in UK Holstein-Friesian heifers. J. Dairy Sci. 94:5413–5423. doi:10.3168/jds.2010-4040.

Eaglen, S.A.E., M.P. Coffey, J.A. Woolliams, and E. Wall. 2012. Evaluating alternate models to estimate genetic parameters of calving traits in United Kingdom Holstein-Friesian dairy cattle. Genet. Sel. Evol. 44:1–13.

Gianola, D. 1982. Theory and analysis of threshold characters. J. Anim. Sci. 54:1079–1096.

Hoeschele, I. 1988. Comparison of “Maximum A-Posteriori Estimation” and “Quasi Best Linear Unbiased Prediction” with threshold characters. J. Anim. Breed. Genet. 105:337–361. doi:10.1111/j.1439-0388.1988.tb00307.x.

Interbull. 2013. Description of National Genetic Evaluations Systems for dairy cattle traits as applied in different Interbull member countries.

López de Maturana, E.L., D. Gianola, G.J.M. Rosa, and K.A. Weigel. 2009. Predictive ability of models for calving difficulty in US Holsteins. J. Anim. Breed. Genet. 126:179–188. doi:10.1111/j.1439-0388.2008.00792.x.

López de Maturana, E.L., A. Legarra, L. Varona, and E. Ugarte. 2007a. Analysis of Fertility and Dystocia in Holsteins Using Recursive Models to Handle Censored and Categorical Data. J. Dairy Sci. 90:2012–2024. doi:10.3168/jds.2005-442.

López de Maturana, E.L., E. Ugarte, and O. González-Recio. 2007b. Impact of Calving Ease on Functional Longevity and Herd Amortization Costs in Basque Holsteins Using Survival Analysis. J. Dairy Sci. 90:4451–4457. doi:10.3168/jds.2006-734.

Luo, M.F., P.J. Boettcher, J.C.M. Dekkers, and L.R. Schaeffer. 1999. Bayesian analysis for estimation of genetic parameters of calving ease and stillbirth for Canadian Holsteins. J. Dairy Sci. 82:1848–e1.

Luo, M.F., P.J. Boettcher, L.R. Schaeffer, and J.C.M. Dekkers. 2001. Bayesian inference for categorical traits with an application to variance component estimation. J. Dairy Sci. 84:694–704.

Matos, C.A., D.L. Thomas, D. Gianola, M. Perez-Enciso, and L.D. Young. 1997. Genetic analysis of discrete reproductive traits in sheep using linear and non-linear models: II. Goodness of fit and predictive ability. J. Anim. Sci. 75:88–94.

Mee, J.F. 2008. Prevalence and risk factors for dystocia in dairy cattle: A review. Vet. J. 176:93–101. doi:10.1016/j.tvjl.2007.12.032.

Misztal, I., D. Gianola, and J.L. Foulley. 1989. Computing aspects of a non-linear method of sire evaluation for categorical data. J. Dairy Sci. 72:1557–1568.

Misztal, I., S. Tsuruta, B. Auvray, and J.M. Lee. 2002. BLUPF90 and related programs (BGF90). Montpellier, France. 743–744.

Olesen, I., M. Perez-Enciso, D. Gianola, and D.L. Thomas. 1994. A comparison of normal and nonnormal mixed models for number of lambs born in Norwegian sheep. J. Anim. Sci. 72:1166–1173.

Phocas, F., and D. Laloë. 2003. Evaluation models and genetic parameters for calving difficulty in beef cattle. J. Anim. Sci. 81:933–938.

Ramirez-Valverde, R., I. Misztal, and J.K. Bertrand. 2001. Comparison of threshold *vs* linear and animal *vs* sire models for predicting direct and maternal genetic effects on calving difficulty in beef cattle. J. Anim. Sci. 79:333–338.

Steinbock, L., A. Näsholm, B. Berglund, K. Johansson, and J. Philipsson. 2003. Genetic effects on stillbirth and calving difficulty in Swedish Holsteins at first and second calving. J. Dairy Sci. 86:2228–2235.

Varona, L., I. Misztal, and J.K. Bertrand. 1999. Threshold-linear versus linear-linear analysis of birth weight and calving ease using an animal model: II. Comparison of models. J. Anim. Sci. 77:2003–2007.

Visscher, P.M., and M.E. Goddard. 1993. Fixed and random contemporary groups. J. Dairy Sci. 76:1444–1454.

Wang, C.S., R.L. Quaas, and E.J. Pollak. 1997. Bayesian analysis of calving ease scores and birth weights. Genet. Sel. Evol. 29:117.

Weller, J.I., and D. Gianola. 1989. Models for genetic analysis of dystocia and calf mortality. J. Dairy Sci. 72:2633–2643.

Weller, J.I., I. Misztal, and D. Gianola. 1988. Genetic analysis of dystocia and calf mortality in Israeli-Holsteins by threshold and linear models. J. Dairy Sci. 71:2491–2501.

Wiggans, G.R., I. Misztal, and C.P. Van Tassell. 2003. Calving Ease (Co)Variance Components for a Sire-Maternal Grandsire Threshold Model. J. Dairy Sci. 86:1845–1848.

CHAPTER III.2 GENETIC EVALUATION OF CALVING EASE FOR WALLOON HOLSTEIN DAIRY CATTLE

Based on: Vanderick, S., T. Troch, A. Gillon, G. Glorieux, P. Faux, and N. Gengler. 2013. Genetic evaluation of calving ease for Walloon Holstein dairy cattle. INTERBULL Bull. 47:32-37.

CHAPTER III.2 GENETIC EVALUATION OF CALVING EASE FOR WALLOON HOLSTEIN DAIRY CATTLE

OUTLINE

Results presented in the Chapter III.1 provided evidence that there was no clear advantage to use a threshold maternal animal model rather than a linear maternal animal model to evaluate calving ease for Walloon Holstein dairy cattle. However, in terms of technically feasibility, the linear model seemed more suitable to implement the routine genetic evaluation of calving ease for the Walloon dairy cattle, thereby enabling to participate to the Interbull international evaluations for this trait. Therefore, this chapter addressed the implementation in routine of this model for the genetic evaluation of calving ease to provide, to Walloon breeders and producers, useful breeding values for direct and maternal effects of calving ease for their Holstein animals in order to help them in their selection and mating decisions.

ABSTRACT

Calving complications have an incidence on the economic profitability of dairy herds. In the Walloon Region of Belgium, calving ease data recording is being done on voluntary basis since 2000. This allows now the implementation of a genetic evaluation of Holstein dairy cattle addressing the need of dairy breeders to select bulls in order to reduce frequency of calving problems. Calving ease scores were analysed using univariate linear animal models, which were fitted with the genetic correlation between direct and maternal additive genetic effects either estimated or constrained to zero. Variance components and related genetic parameters were estimated from a data set including 33 155 calving records. Included in the models were fixed season effects, fixed herd effects and fixed sex of calf × age of dam classes × group of calvings interaction effects, random herd × year of calving effects, random maternal permanent environment effects, and random animal direct and maternal additive genetic effects. For both models, direct and maternal heritabilities for calving ease were about 8 and 2%, respectively. Genetic correlation between direct and maternal additive effects was found to be non-significantly different from zero. So, a linear animal model with genetic correlation between direct and maternal effects constrained to zero was adopted for the routine genetic evaluation of calving ease for Walloon Holstein dairy cattle. This model was validated by Interbull in January 2013 and, since April 2013; the Walloon Region of Belgium has officially participated to the international MACE evaluation for calving traits.

Keywords: calving ease, animal model, Holstein

INTRODUCTION

Calving complications impact production, fertility and cow and calf morbidity and mortality and affect negatively economic profitability in dairy herds (Dematawena and Berger, 1997; Eaglen *et al.*, 2011). Calving ease measures the presence or absence of calving problems and their intensities. This trait is generally scored on a categorical scale by the breeder, what makes it more sensitive to subjectivity (Dekkers, 1994).

Additionally, calving ease is affected by two additive genetic components, the ability of the calf to be easily born (direct effect) and the ability of the dam to easily give birth (maternal effect). The direct effect is expressed only once, when the calf is born whereas the maternal effect is expressed several times, each time a cow calves.

Considering the categorical nature of the trait, from a theoretical point of view, application of a threshold model is the correct choice (Gianola, 1982), whereas, from a practical point of view, linear model is a more easily applicable choice (Varona *et al.*, 1999; Ramirez-Valverde *et al.*, 2001; Phocas and Laloë, 2003). This is evidenced by the fact that different Interbull members use a linear approach for the genetic evaluation of categorical calving traits (Interbull, 2013).

The aim of this research was to estimate the genetic parameters for direct and maternal additive effects based on a linear animal model and to report the implementation of a routine genetic evaluation for calving ease in the Walloon Holstein dairy cattle.

MATERIALS AND METHODS

DATA MATERIAL FOR CALVING EVALUATION

Calving ease is scored by dairy breeders on voluntary basis and collected by the Walloon Breeding Association (AWE). Calving ease scores ranged from 1 to 4 (1=caesarean and embryotomy, 2=hard pull, 3=easy pull, and 4=normal). The original data set comprised 138 144 calving records and presented a typical distribution of CE; most of the records fell into category 4 (69%) and few records into category 1 (about 1%).

Records from Holstein calves born between 2000 and 2012 were used. Edits were done for out-of-range values for calving ease and records with missing information related to the factors in the statistical model, were removed. Multiple births and calf without dam known were excluded. Records were limited to first five parities. Age at calving must be 21-48 months for heifers (1st parity) and 31-142 months for cows (2nd-5th parities) with specific limits depending on parity. Since dairy breeders use their own judgment to assign scores for calving ease, data quality depends highly on them. Therefore, herds with a standard deviation for scores smaller than 0.05 were deleted in order to avoid herds where breeders put all scores in the same category. Herds had to report at least, in average, four calvings per year calculated on the first two parities. The full data set contained calving ease records from 85 118 calves born from 62 265 dams in 862 herds. Pedigree data were extracted from the database used for the official Walloon genetic evaluations. The final pedigree file included 233 882 animals. More details in Table 5.

ESTIMATION OF (CO)VARIANCE COMPONENTS AND RELATED PARAMETERS

To ensure an appropriate data structure for the parameter estimation, some extra edits were applied. Calves were required to have known sire and dams were required to display a calving ease record in 1st parity. Only data from continuous calvings per dam were kept (*e.g.* if a dam displayed calving ease records from its first, second and fourth calving only calving ease records from first and second calving were kept). Herds had to display at least, in average, four calvings per year calculated on the 1st parity only. A final edit required in average more than one calving per dam per herd in order to avoid the presence of herds where only heifers were recorded. Based on these edits, in order to estimate (co)variance components, calving ease records from 33 155 calves born from 25 240 dams and 2215 sires in 492 Walloon herds were used. The final pedigree contained 120 374 animals. Table 5 displays characteristics of this data set.

Table 5. Characteristics of full and variance component estimation (VCE) data sets

	Full data set		VCE data set	
	N	%	N	%
Calving ease records	85 118		33155	
Female calves	66 511	78.1	26 177	78.9
Male calves	18 511	21.9	6978	21.1
Herds	862		492	
Sires with progeny records	3148		2215	
Dams	62 265		25 240	
Final pedigree file	233 882		120 374	
Calving ease scores				
1. Caesarean and embryotomy	781	0.9	443	1.3
2. Hard pull	4006	4.7	2179	6.6
3. Easy pull	23 461	27.6	10 114	30.5
4. Normal	56 870	66.8	20 419	61.6

MODELS FOR CALVING EASE EVALUATION

Preliminary analyses were carried out using general linear model procedure of SAS. Different linear models were fitted aiming to identify fixed effects that affected variation of calving scores.

The ultimate model included the fixed effects of season (4 classes), herd and sex of calf × age of dam classes (11 classes) × group of parities (2 classes: 1st parity and 2nd-5th parities). These effects were included in the following univariate linear animal model; where calving ease was modelled as continuous trait; to estimate (co)variance components:

$$\mathbf{y}_{\mathrm{CE}} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_h\mathbf{h} + \mathbf{Z}_a\mathbf{a} + \mathbf{Z}_m\mathbf{m} + \mathbf{Z}_p\mathbf{p} + \mathbf{e} \qquad (44)$$

where $\boldsymbol{\beta}$ is a vector of fixed effects, $\mathbf{h}$ is a vector of random herd$\times$year of calving effects, $\mathbf{a}$ is a vector of random direct additive genetic effects, $\mathbf{m}$ is a vector of random maternal additive genetic effects, $\mathbf{p}$ is a vector of random permanent maternal environmental effects, $\mathbf{X}$, $\mathbf{Z}_h$, $\mathbf{Z}_a$, $\mathbf{Z}_m$ and $\mathbf{Z}_p$ are known incidences matrices linking data with respective effects, and $\mathbf{e}$ are the residuals. This model was fitted with the genetic correlation between direct and maternal additive genetic effects either estimated or constrained to null; called Model L1 and Model L2, respectively, in the remainder of this paper.

Procedures based on a Bayesian approach using the Gibbs sampling algorithm were used to estimates (co)variance components and related genetic parameters by using programs gratefully provided by Ignacy Misztal (Misztal *et al.*, 2002). No genetic grouping was used.

GENETIC EVALUATION SYSTEM

For routine genetic evaluation of calving ease, the univariate linear animal model (44) was applied on the full data set and solved by sparse inversion. Estimated breeding values (EBV) of calving ease were expressed in expected differences in percentage of “normal” calving. Reliability (REL) of EBV was defined as the squared correlation between true and predicted EBV. It was estimated based on the diagonal elements of the mixed model equations, as shown by Henderson (1984). Prediction Error Variance (PEV) given by direct inversion of the coefficient matrix allowed the estimation of correct reliabilities:

$$REL_x = \frac{\sigma_x^2 - PEV_x}{\sigma_x^2} \qquad (45)$$

where σ_x^2 is the additive genetic variance for genetic component x (*i.e.* direct or maternal effect).

RESULTS AND DISCUSSION

ESTIMATES OF (CO)VARIANCE COMPONENTS AND OF RELATED GENETIC PARAMETERS

Results for the (co)variance components, related genetic parameters, and fraction part of variances in phenotypic variance for Model L1 and Model L2, are reported in Table 6. Parameters can generally be considered as significantly different from zero, posterior means being greater than twice the posterior standard deviation, except for genetic correlation between direct and maternal additive genetic effects for Model L1.

The additive genetic variance due to direct effects was greater than that due to maternal effects for both models. Indeed, heritabilities for direct effects were nearly four times higher than for maternal effects. Heritabilities estimates were in the range of previously published estimates of this trait in dairy cattle, which ranged from 0.03 to 0.17 for direct heritability (h^2) and from 0.02 to 0.12 for maternal h^2 (Weller and Gianola, 1989; Steinbock *et al.*, 2003; López de Maturana *et al.*, 2007; Eaglen *et al.*, 2012).

The herd × year of calving effects represented 12% of the phenotypic variance, which was the largest contributor to the phenotypic variance after the residual effects (74%). Effects of herd × year of calving take into account biological differences among herds and years of calving but also subjectivity of calving ease scores.

The maternal permanent environment effects represented 5% of the phenotypic variance in each model and were greater than the genetic maternal effects. Preliminary analyses to this study have established that maternal genetic variance tended to be overestimated by models taking not into account the existence of a maternal permanent environment effects. Hence, most of the observed variability due to the maternal effects was found to be of a non (additive) genetic origin. As a special care was taken to get useful data for variance component estimation (*e.g.* maximizing repeated calvings), this result should not be an artefact of the sparseness of repeated records used.

Table 6. Posterior mean (PM) and posterior standard deviation (PSD) of (co)variance components and related genetic parameters

	Model L1		Model L2	
Parameter[a]	PM	PSD	PM	PSD
σ_h^2	0.042	0.002	0.042	0.002
σ_a^2	0.027	0.004	0.028	0.004
σ_m^2	0.008	0.003	0.009	0.002
σ_p^2	0.018	0.004	0.017	0.005
σ_e^2	0.269	0.005	0.269	0.005
$r_u(a,m)$	0.088	0.194	-	-
h_a^2	0.074	0.012	0.078	0.012
h_m^2	0.023	0.007	0.024	0.007
C_h	12%		12%	
C_p	5%		5%	
C_e	74%		74%	

[a] The terms σ_h^2 is the herd × year of calving variance, σ_a^2 is the direct additive genetic variance, σ_{m}^2 is the maternal additive genetic variance, σ_p^2 is the permanent maternal environmental variance, σ_e^2 is the residual variance, $r_u(a,m)$ is the genetic correlation between direct and maternal effects. C_h, C_p and C_e are the herd × year of calving fraction, permanent maternal environmental fraction and residual fraction in the phenotypic variance, respectively.

Estimates of variance components and related genetic parameters were similar between Model L1 and Model L2; though genetic correlation between direct and maternal additive genetic effects was estimated or constrained to zero, indicating that the value of this correlation had no large influence. Although a low positive genetic correlation was estimated between genetic effects with Model L1, it seemed more relevant to consider no genetic correlation according to its posterior standard deviation.

GENETIC EVALUATION RESULTS

According to previous statements, Model L2 was applied on the full data set. Figure 1 shows genetic trends of direct and maternal EBV of Holstein AI bulls. Those bulls were required to have a minimal REL of 0.35.

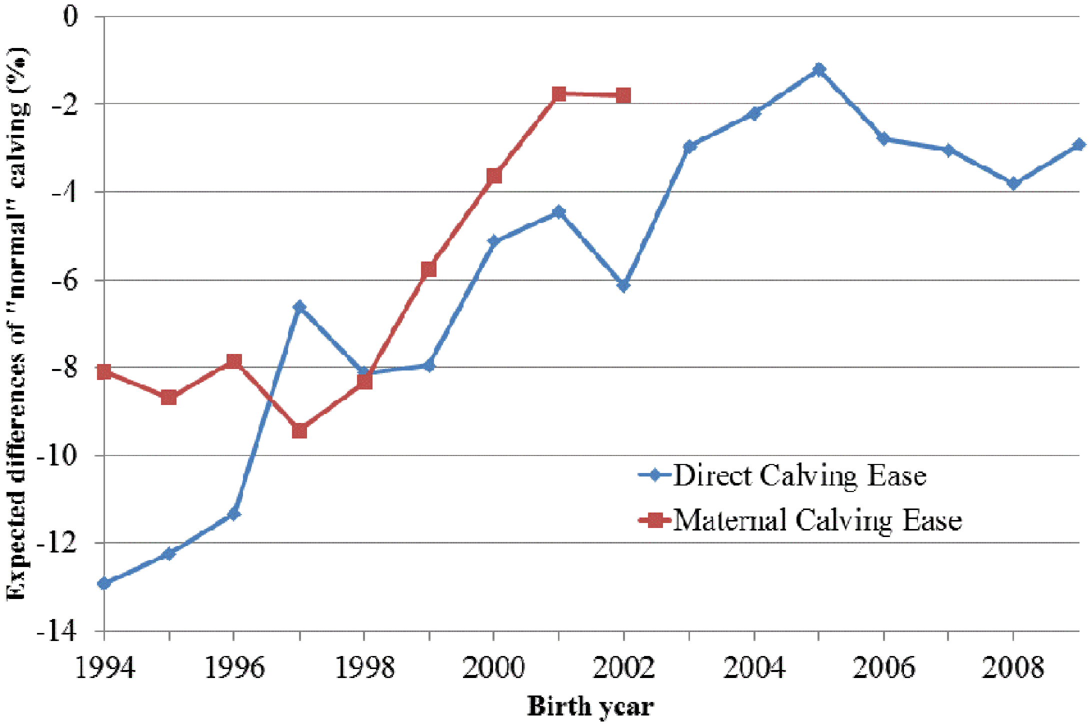


Figure 1. Genetic trends of direct and maternal breeding values of calving ease in Holstein bulls with a minimal REL of 0.35

In Figure 1, the number of bulls was 1171 and 202 for direct and maternal EBV, respectively. Mean, standard deviation, minimum and maximum of these bulls' EBV are shown in Table 7.

Few bulls did meet the REL requirement for maternal genetic effect and so there were not enough bulls to provide reasonable results after a birth year of 2002.

Table 7. Descriptive statistics of direct and maternal calving ease breeding values for bulls from trends

	Mean	SD	Min	Max
Direct calving ease	-5.11	9.21	-40.09	30.66
Maternal calving ease	-6.55	8.64	-31.10	18.47

There were clear positive genetic trends for both effects although there was no genetic evaluation system of calving ease for Walloon dairy cattle. As currently the testing of young sires is extremely limited in the Walloon Region of Belgium this reflects the preselection done in exporting countries before these sires are used.

Figure 2 shows that similar genetic trends were observed in Holstein cows and calves. Animals used to realise these trends were required to have a minimal REL of 0.15. This can be explained by the fact that foreign scale calving information for used sires was known before and most likely used by breeders. Even if this information was suboptimal it allowed avoiding sires with major calving problems.

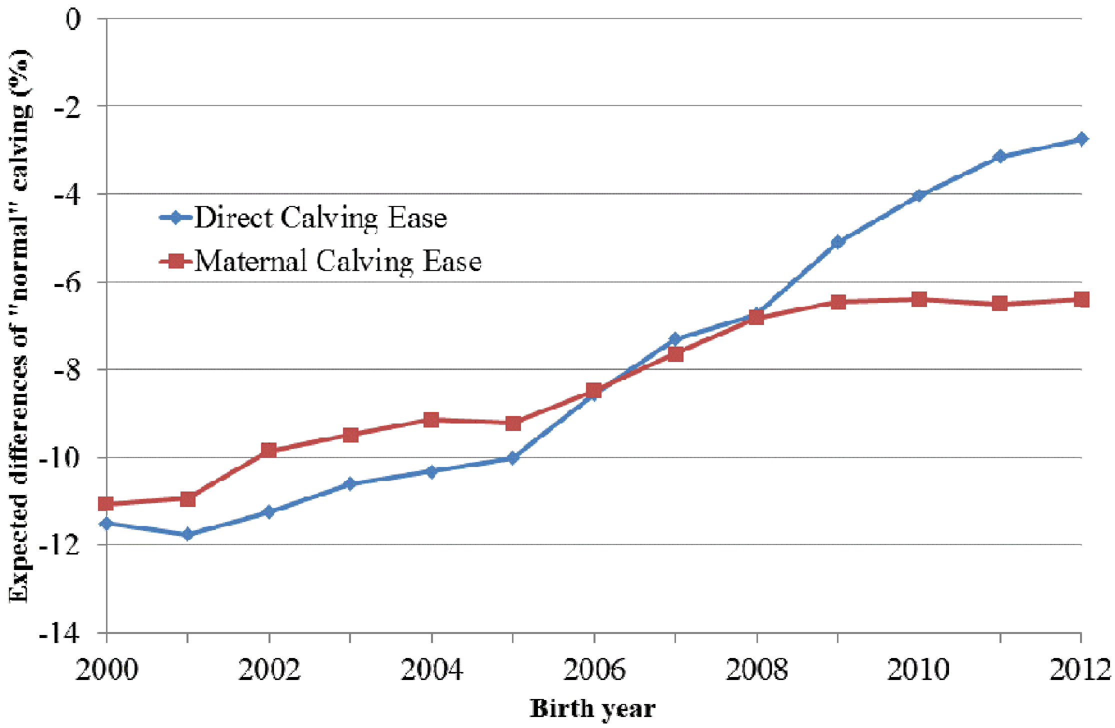


Figure 2. Genetic trends of direct and maternal breeding values for calving ease in Holstein cows and calves with a minimal REL of 0.15

Based on these results using Model L2, Region Walloon of Belgium participated in Interbull January 2013 MACE test run for Holstein breed. Model L2 was validated with Interbull trend validation method III. Genetic correlations with other countries were estimated for direct and maternal genetic effects of CE. On average, country correlation was 0.62 for direct calving ease and 0.64 for maternal calving ease, respectively. The highest country correlation was with Switzerland Red and White Holstein (0.73) for direct calving ease and with Germany (0.75) for maternal calving ease, respectively.

CONCLUSIONS

A univariate linear animal model was developed for routine genetic evaluation of calving ease for Walloon Holstein dairy cattle. No relevant genetic correlation between direct and maternal effects on ease of calving was found. Model L2 was therefore implemented in routine in April 2013 to provide, to Walloon dairy breeders and producers, EBV of calving ease for Holstein cattle.

ACKNOWLEDGEMENTS

The Ministry of Agriculture of the Walloon Region of Belgium (Service Public de Wallonie, Direction générale opérationnelle “Agriculture, Ressources naturelles et Environnement) is acknowledged for its financial support. The authors gratefully acknowledge the technical support provided by the Walloon Breeding Association (Ciney, Belgium).

LITERATURE CITED

Dekkers, J.C.M. 1994. Optimal breeding strategies for calving ease. J. Dairy Sci. 77:3441–3453.

Dematawena, C.M.B., and P.J. Berger. 1997. Effect of dystocia on yield, fertility, and cow losses and an economic evaluation of dystocia scores for Holsteins. J. Dairy Sci. 80:754–761.

Eaglen, S.A.E., M.P. Coffey, J.A. Woolliams, R. Mrode, and E. Wall. 2011. Phenotypic effects of calving ease on the subsequent fertility and milk production of dam and calf in UK Holstein-Friesian heifers. J. Dairy Sci. 94:5413–5423. doi:10.3168/jds.2010-4040.

Eaglen, S.A.E., M.P. Coffey, J.A. Woolliams, and E. Wall. 2012. Evaluating alternate models to estimate genetic parameters of calving traits in United Kingdom Holstein-Friesian dairy cattle. Genet. Sel. Evol. 44:1–13.

Gianola, D. 1982. Theory and analysis of threshold characters. J. Anim. Sci. 54:1079–1096.

Henderson, C.R. 1984. Applications of linear models in animal breeding. 3th ed. University of Guelph, [Guelph, Ont.]. 384 pp.

Interbull. 2013. Description of National Genetic Evaluations Systems for dairy cattle traits as applied in different Interbull member countries.

López de Maturana, E.L., A. Legarra, L. Varona, and E. Ugarte. 2007. Analysis of Fertility and Dystocia in Holsteins Using Recursive Models to Handle Censored and Categorical Data. J. Dairy Sci. 90:2012–2024. doi:10.3168/jds.2005-442.

Misztal, I., S. Tsuruta, B. Auvray, and J.M. Lee. 2002. BLUPF90 and related programs (BGF90). Montpellier, France. 1–2.

Phocas, F., and D. Laloë. 2003. Evaluation models and genetic parameters for calving difficulty in beef cattle. J. Anim. Sci. 81:933–938.

Ramirez-Valverde, R., I. Misztal, and J.K. Bertrand. 2001. Comparison of threshold *vs* linear and animal *vs* sire models for predicting direct and maternal genetic effects on calving difficulty in beef cattle. J. Anim. Sci. 79:333–338.

Steinbock, L., A. Näsholm, B. Berglund, K. Johansson, and J. Philipsson. 2003. Genetic effects on stillbirth and calving difficulty in Swedish Holsteins at first and second calving. *J. Dairy Sci.* 86:2228–2235.

Varona, L., I. Misztal, and J.K. Bertrand. 1999. Threshold-linear versus linear-linear analysis of birth weight and calving ease using an animal model: II. Comparison of models. *J. Anim. Sci.* 77:2003–2007.

Weller, J.I., and D. Gianola. 1989. Models for genetic analysis of dystocia and calf mortality. *J. Dairy Sci.* 72:2633–2643.

CHAPTER III.3 DERIVATION OF A NEW LAMB SURVIVAL TRAIT FOR THE NEW ZEALAND SHEEP INDUSTRY

Based on: Vanderick, S., B. Auvray, S.-A. Newman, K. G. Dodds, N. Gengler, and J. M. Everett-Hincks. 2015. Derivation of a new survival trait for the New Zealand sheep industry. J. Anim. Sci. 93:3765-3672.

CHAPTER III.3 DERIVATION OF A NEW LAMB SURVIVAL TRAIT FOR THE NEW ZEALAND SHEEP INDUSTRY

OUTLINE

The second maternally influenced trait studied was lamb survival to weaning using data provided by the New Zealand sheep industry. The genetic improvement of this trait has become increasingly important because mortality in lambs causes substantial economic loss as well as animal welfare concerns. A genetic evaluation system of this trait existed already; however, previous research indicated the need to review the genetic evaluation model to increase its accuracy and, thus, for improving the rate of genetic gain of lamb survival to weaning to reduce lamb mortality rate. Therefore, the main objective of this chapter was to examine different maternal animal models to analyse lamb survival data. The benefits to apply or not a logit transformation of the lamb survival data were also examined due to the discrete nature of this trait. The (co)variance components and the derived genetic parameters for lamb survival to weaning were estimated using different linear and logistic maternal animal models. Then, maternal effects and genetic correlations were investigated and models were assessed in terms of predictive ability. Another issue addressed in this chapter was to define new rules for data inclusion and exclusion in the genetic analysis of lamb survival to weaning. Data used in this chapter were provided by Sheep Improvement Limited (SIL), Christchurch, New Zealand.

ABSTRACT

Previous research identified that a review of the current industry New Zealand lamb survival trait was necessary as its recording accuracy was reliant on farmers notifying their Sheep Improvement Limited bureau of lamb deaths. This paper reports the decision rules and genetic parameters for a new lamb survival trait for the New Zealand sheep industry. These rules define the new lamb survival trait (NEWSUR) using lamb birth fate (BFATE) codes and the presence/absence of lamb weight measurements. Six univariate animal models were tested and used to estimate variance or covariance components and the resulting direct and maternal heritabilities for NEWSUR. The models differed in the way they adjust for the effect of day of birth, the exclusion or inclusion of a litter (dam/year of birth) random effect and the application or not of a logit transformation of the phenotypes. For both the linear and logistic methods, models including the random effect of litter provided the best fit for NEWSUR according to log-likelihood values. Log-likelihoods for the linear and logistic models cannot be compared, therefore a cross-validation method was used to assess whether the logit transformation was appropriate by analysing the predictive ability of the models. The mean square errors were slightly lower for the linear compared to the logistic model and therefore the linear model was recommended for industry use. The heritability (h^2) attributed to direct effects ranged from 2 to 5.5%. A direct h^2 of 5.5% resulted from a linear model without litter effect and omitting the effect of day of birth on survival, whereas a direct h^2 of 2% resulted from the logistic model fitting a random litter effect. The h^2 attributed to maternal genetic effects ranged from 1.9 to 7.7%. A maternal genetic h^2 of 7.7% resulted from the logistic model omitting the litter effect, whereas a maternal genetic h^2 of 1.9% resulted from the linear model fitting a random litter effect. The addition of the litter random effect decreased the maternal heritabilities substantially in all cases and was recommended for industry use to avoid overestimation of the maternal genetic variance. SIL has implemented NEWSUR and the associated genetic evaluation model based on information described in this paper. Industry-wide implementation will enable sheep breeders to produce more accurate genetic evaluations to their commercial clients.

Keywords: lamb survival, maternal genetics, litter effect, heritability.

INTRODUCTION

Lambing percentage is one of the most significant traits affecting profitability on New Zealand sheep farms (Hawkins and Wu, 2011). Since the early 1990s, lambing percentage has increased at about 1% per year from a relatively stable level of approximately 100% (McDermott *et al.*, 2008). As lambing percentage increases, the proportion of ewes having twin and triplet litters increases. Lamb mortality rate in multiples is higher than in singles, with triplets being particularly susceptible (Everett-Hincks and Dodds, 2008). Improving lamb survival has become increasingly important to the New Zealand sheep industry.

Lamb survival is a complex trait influenced by direct genetic, maternal genetic and environmental effects (Brien *et al.,* 2014). Low h^2 estimates have been reported across a number of studies meaning that lamb survival is predominantly controlled by environmental factors, limiting its genetic improvement (Amer and Jopson, 2003; Safari *et al.*, 2005).

Sheep Improvement Ltd (SIL; http://www.sil.co.nz), is New Zealand's national sheep genetic evaluation system and records lamb survival to weaning but genetic progress has been limited due to the low h^2 of the trait and also due to the current method of farmer recording.

The objectives of this paper are 3-fold: first, to review the current SIL lamb survival trait (SILSUR) and develop a more accurate and consistently recorded lamb survival trait (new lamb survival trait [NEWSUR**]**) for industry implementation; second, to define new decision rules for data inclusion and exclusion in SIL genetic evaluation system; and lastly, to revise the genetic evaluation model, estimate variance components and investigate maternal genetic effects and correlations to produce estimated breeding values (EBV) for NEWSUR.

This research will allow SIL to adopt an improved genetic evaluation for lamb survival, which will in turn increase the rate of genetic gain through the design of appropriate animal breeding programs to reduce lamb mortality rates.

MATERIALS AND METHODS

DATA

Animal performance records were extracted from the SIL national sheep recording database in October 2011 and consisted of 3 427 496 lamb records, born between 1990 and 2010 from 596 performance recorded flocks. These flocks cover a range of breeds and crosses (2 breed combinations), but are predominantly Romney, Coopworth, Perendale, Texel, and some composites (defined as a combination of 3 or more breeds; Price, 2000).

LAMB SURVIVAL TRAIT DEFINITIONS

Lamb survival is recorded as a binary trait in SIL. Lambs that survive to weaning are given a "1" and those that die are given a "0". However the SIL system relies on notification from the sheep breeder that a lamb has died by assigning a birth fate (BFATE) code (Table 8); otherwise, they are assumed alive at weaning.

Table 8. Sheep Improvement Limited lamb birth fate (BFATE) codes

BFATE code	Description
E	Embryo transfer progeny
F	Fostered
H	Hand-reared
J	Born dead
J3	Died within 3 days of birth (autopsy)
K	Died between birth and rearing
L	AI progeny
M	Died of misadventure
P	Born dead - premature (autopsy)
R	Born dead - rotten (autopsy)
1	Died between rearing and weaning
4	Culled at birth (alive but not tagged)

Rules to define SILSUR are described in Figure 3; lambs without BFATE record are considered alive and lambs with a BFATE record are considered dead except for lambs with BFATE "E" (embryo transfer progeny), "L" (AI progeny), "M" (died of misadventure) or "4" (culled at birth). Lambs with BFATE "E" or BFATE "L" are treated similarly to lambs without BFATE, because those birth fates have no effect on genetic merit. Survival of lambs with BFATE "M" are set to missing so they are removed from the analysis and lambs with BFATE "4" are lambs alive but not tagged.

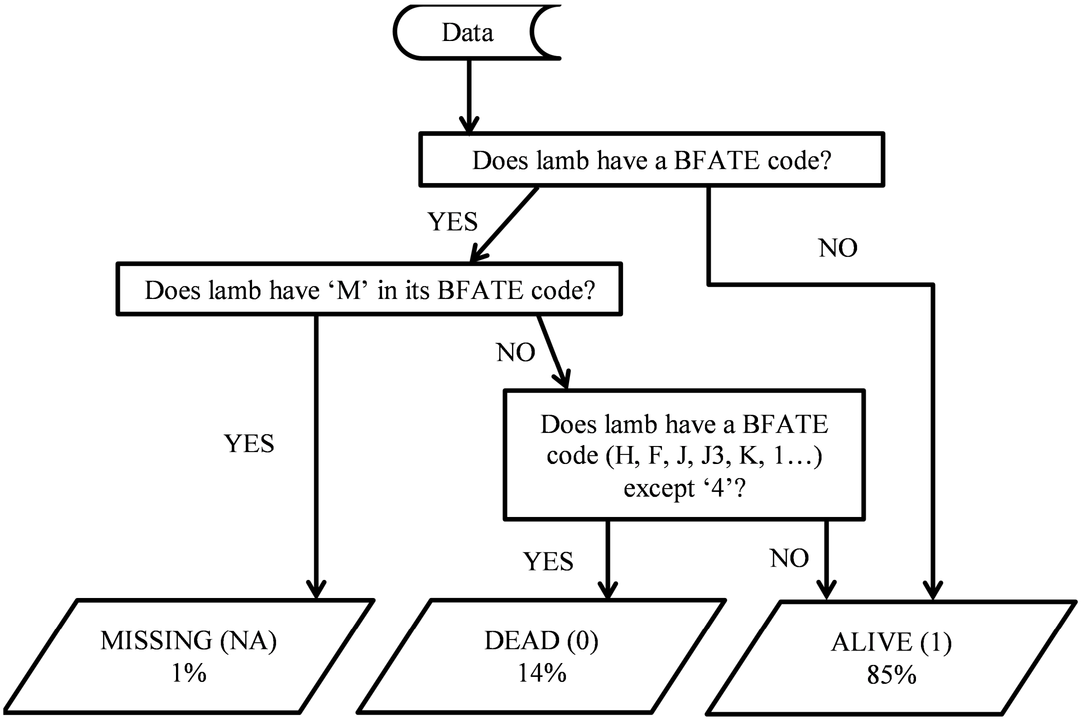


Figure 3. The current Sheep Improvement Limited lamb survival trait. BFATE= lamb birth fate; M = died of misadventure; H = hand reared; F = fostered; J = born dead; J3 = died within 3 days of birth (autopsy); K = died between birth and rearing; 1 = died between rearing and weaning; 4 = culled at birth (alive but not tagged); NA = not available.

Besides, SILSUR data rules exclude records where the mean lamb survival for a flock and year of birth is <55 or ≥93% as it is assumed that farmer recording errors have largely led to survival rates below and above these limits. (These will be referred to as "survival limits" throughout the remainder of this document). Moreover, some flocks have had their data permanently excluded for particular years with known issues. Proportions of survival reported in Figure 3 were observed after application of the survival limits.

The definition of NEWSUR is based on using a combination of BFATE codes and the presence/absence of weight measurements to assess whether a lamb has survived until weaning age (*i.e.* 100 days). The weight measurements used are weaning weight, live weight at 6 months of age, live weight at 8 months of age, live weight at 10 months of age, live weight at 12 months of age, live weight at 18 months of age and greasy fleece weight at 12 months of age. The rules used to assign NEWSUR are described in Figure 4. Lambs with BFATE “E” (embryo transfer progeny) or “L” (AI progeny) are treated similarly to lambs without BFATE. Then, if these lambs have at least 2 recorded weights, they will be considered alive; otherwise their survival will be set to missing. Lambs with BFATE “H” are counted as dead because lambs would have died without breeder’s intervention. Lambs with BFATE “F” are considered missing because some breeders foster triplets regularly as a standard management practice for their flock (54% of lambs with a BFATE code containing “F” have a birth rank of 3). The NEWSUR uses the same survival limits as those used for SILSUR. Proportions of survival reported in Figure 4 were observed after application of the survival limits.

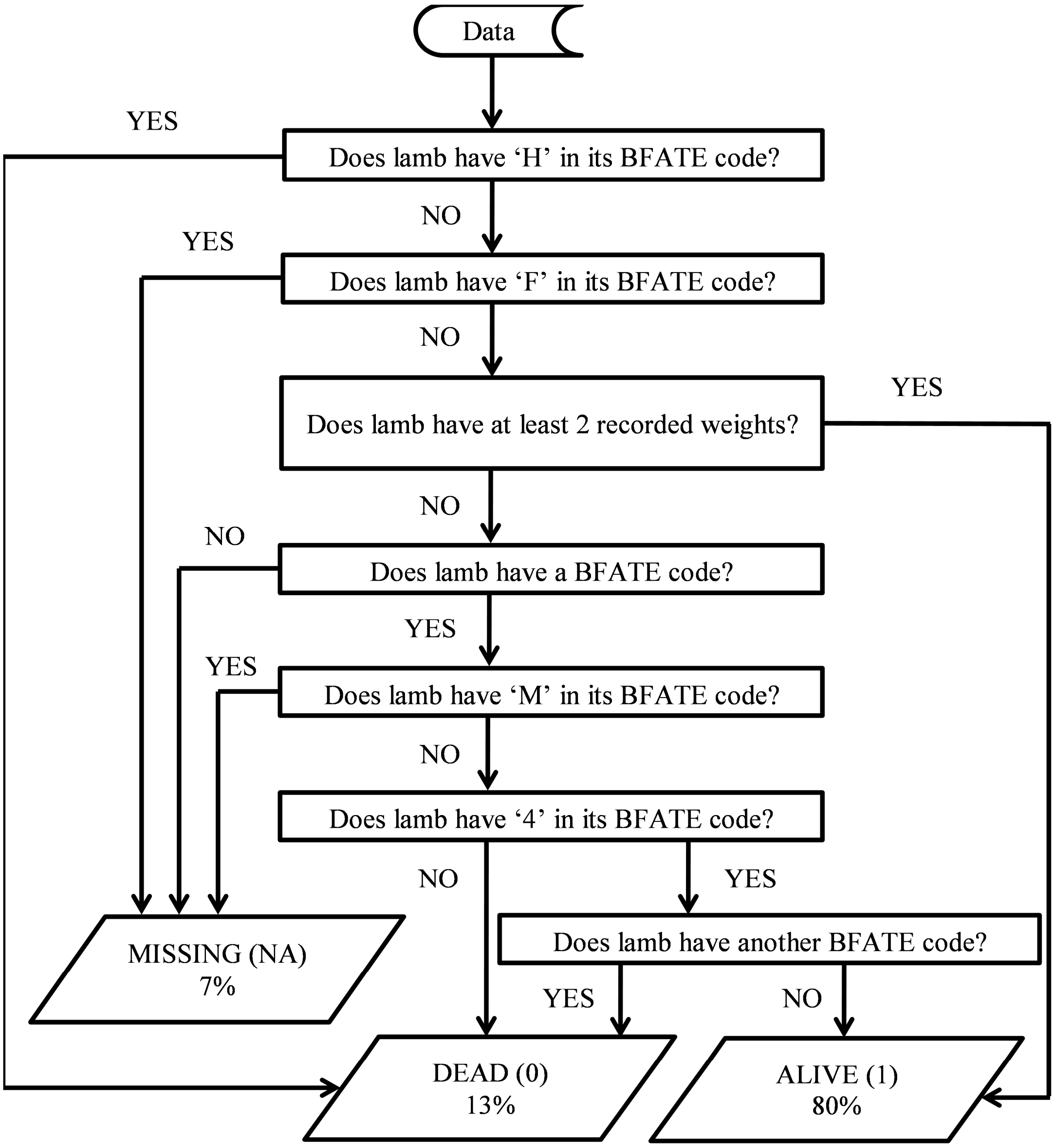


Figure 4. Rules used to define the new lamb survival trait. BFATE = lamb birth fate; H = hand reared; F = fostered; M = died of misadventure; 4 = culled at birth (alive but not tagged); NA = not available.

ESTIMATION OF BREEDING VALUES FOR THE NEW LAMB SURVIVAL TRAIT

Data Preparation

The NEWSUR data were edited to remove missing records and effect levels with low number of observations after applying SIL survival limits, resulting in data sets of differing sizes for NEWSUR and SILSUR. For age of dam classes (AOD), hogget dams (1-year-old ewes) or dams with missing age or age <1 year old were removed. Dams older than 7 years of age were grouped into age group 7. Lambs from litter sizes greater than 3 (birth rank > 3) were removed from the data set. Lambs with a missing birth date were also removed. Sex is known to affect lamb survival but the industry standard is to not record the sex of lambs found dead at tagging. Including 3 classes of sex (male, female, and unknown) in the analysis was not appropriate as unknown sex animals had zero survival. Accordingly, for the purpose of estimating variance components, the sex of dead lambs was assigned at random with equal probability for each sex as done by Lopez-Villalobos and Garrick (1999). To estimate maternal genetic effects accurately, at least 3 generations of animals were required and, therefore, 6 year of records were included in the analysis. A subset of 6 years of lambing data (2005 to 2010) was created for NEWSUR and then used to estimate the variance components of NEWSUR. In this subset, the 944 211 lambs (from 212 flocks) were the progeny of 16 084 sires and 353 783 dams with 5 418 paternal grandsires, 10 798 paternal grandmothers, 13 590 maternal grandsires and 179 359 maternal grandmothers. There were 144 636 single-born lambs, 647 008 twin-born lambs and 152 567 triplet-born lambs of various breeds and breed composition (predominantly Romney, Perendale, Coopworth, and Texel).

Genetic Parameter Estimation

All analyses were performed using ASReml 3 (Gilmour *et al.*, 2009). Six univariate animal models were tested to estimate variance or covariance components and genetic parameters for NEWSUR were derived. All models included the fixed effects of flock (212 flocks), year of birth of the lamb (6 levels: 2005 to 2010), litter size at birth or birth rank (3 levels: single, twin, and triplet), AOD in years (6 levels: 2-7 year) and sex of lamb (2 levels: ram and ewe) and their interactions. The random direct and maternal genetic effects and the covariance between direct and maternal genetic effects were also fitted in all models. The different models investigated are summarized in Table 9. Models A1, A2, A3 and B analysed NEWSUR as a normally distributed trait and models LogA and LogB analysed NEWSUR as a logit transformed trait using model A3 and model B, respectively. Models A1, A2 and A3 differ in the way they deal with the explanatory variable “lamb day of birth”. Model A1 ignores lamb day of birth, Model A2 fits day of birth (day 1 to day 355) within flock and birth year as a covariate, and Model A3 fits day of birth classified in periods of 2 weeks (25 classes) within flock and birth year as random effects. Model B is the same as model A3 but has an extra random non-genetic effect that is specific to a litter (interaction between dam and year of birth) fitted. This litter effect reflects common environmental effects unique to the litter as well as non-genetic effects expressed in full sibs and genotype × year interaction effects expressed in the dam.

Heritabilities and genetic correlations were calculated from the variance estimates produced by ASReml 3.0 (Gilmour *et al.*, 2009). The total h^2 was obtained by summing genetic variances attributed to direct and maternal effects and the covariance (doubled). Genetic correlations between direct and maternal genetic effects were calculated by dividing the covariance between direct and maternal effects by the standard deviation of direct and maternal genetic effects.

Table 9. Description of the six models used to estimate variance components

		Model[b]					
Type	Effect[a]	A1	A2	A3	B	LogA	LogB
Fixed class effect	Flk	√	√	√	√	√	√
	BY	√	√	√	√	√	√
	BR	√	√	√	√	√	√
	AOD	√	√	√	√	√	√
	Sex	√	√	√	√	√	√
	Flk × BY	√	√	√	√	√	√
	Flk × BR	√	√	√	√	√	√
	BY × BR	√	√	√	√	√	√
	BR × AOD	√	√	√	√	√	√
	Flk × BY × BR	√	√	√	√	√	√
Fixed covariates	BDAY × Flk × BY		√				
Random effect	direct genetic	√	√	√	√	√	√
	maternal genetic	√	√	√	√	√	√
	cov(a, m)	√	√	√	√	√	√
	BDAYc × Flk × BY			√	√	√	√
	litter				√		√
Phenotypic transformation	logit					√	√

[a] Flk = flock; BY = year of birth class; BR = birth rank class; AOD = age of dam class; BDAY = day of birth; cov(a,m) = covariance between direct and maternal genetic effects; BDAYc = day of birth classified in periods of 2 wk. Litter is the random litter effect (dam/year) and logit is the logit transformation of phenotype.

[b] A1 = model with no day of birth fitted; A2 = model where BDAY × Flk × BY is fitted as fixed covariates; A3 = model where BDAYc × Flk × BY is fitted as random effects; B = same as A3 but with an extra random litter effect fitted; LogA = model A3 but with phenotypes transformed using a logit transformation; LogB = model B with logit transformation.

Comparison of Models Based on their Predictive Ability

First, the best linear model was chosen as the one with the highest restricted likelihood. The same method was used for comparing logistic models. Second, because log-likelihoods for the linear and logistic models cannot be compared, the following procedure was performed to assess whether a logit transformation was appropriate for NEWSUR by analysing the ability of the 2 models to predict the lamb survival phenotype. A 3-fold cross-validation method was used, where survival records from lambs born in 2010 were randomly allocated to 1 of 3 groups, and EBV were estimated 3 times using both models, sequentially setting all observations in 1 of the 3 groups to missing. The phenotypes were then predicted from the 2 fitted models for the subset where the phenotypes had been set to missing. The variances of the difference between observed and predicted values, that is, mean square error (MSE), obtained with both models were then calculated and compared with an intercept model (using the average lamb survival as a predictor).

RESULTS AND DISCUSSION

COMPARISON BETWEEN EXISTING AND NEW SURVIVAL PHENOTYPES

The incidences of NEWSUR and SILSUR, before application of the survival limits, were compared in Table 10. With NEWSUR, 11.3% of lambs recorded alive with SILSUR are now considered missing. Most of these lambs are lambs without a BFATE code and are assigned alive by default by SIL. With NEWSUR, these lambs without a BFATE code and without at least 2 weights are considered missing. Another significant difference between NEWSUR and SILSUR is that 42% of lambs considered missing with SILSUR are now considered alive with NEWSUR. Most of these lambs are lambs with an ‘‘M’’ (died by misadventure) BFATE code (or BFATE code combination, as lambs may have more than one code to describe their fate) and are considered missing by SIL, despite having at least 2 recorded weights.

Table 10. Cross-frequencies of the new lamb survival trait (NEWSUR) and the current Sheep Improvement Limited lamb survival trait (SILSUR; and conditional percentages, per row/column if appropriate)

	SILSUR status			
NEWSUR status	Dead	Alive	Missing	Total
Dead	388 993 (99.6% / 96.3%)	10 901 (0.4%/ 2.7%)	4187 (9.4% / 1.0%)	404 081 (11.8%)
Alive	701 (0.2% / 0.0%)	2 644 398 (88.4% / 99.3%)	18 784 (42.0%/ 0.7%)	2 663 883 (77.7%)
Missing	840 (0.2%/ 0.2%)	336 980 (11.3% / 93.7%)	21 712 (48.6% / 6.0%)	359 532 (10.5%)
Total	390 534 (11.4%)	2 992 279 (87.3%)	44 683 (1.3%)	3 427 496

GENETIC PARAMETERS ESTIMATION

All fixed effects fitted in the 6 models investigated for NEWSUR were significant at $P<0.001$. Estimates of variance or covariance components and derived genetic parameters for NEWSUR are reported in Table 11 for the linear models tested and in Table 12 for the logistic models. Variance components of both logistic models cannot be compared with those obtained for the linear models. However, heritabilities and correlations can be compared as they are ratios of variances.

Estimates of residual variances, direct genetic variances and total phenotypic variances were relatively constant for the linear models (*i.e.* models A1, A2, A3, and B).

Estimates of maternal genetic variances were relatively constant but decreased with the addition of the random effect of litter fitted in model B. This decrease was also observed for the logistic models, indicating that a part of the litter variance has been attributed to maternal genetic variance when the litter random effect was not fitted. Consequently, the maternal genetic variance would be overestimated in models without the litter random effect fitted. This finding is similar to that made by other authors investigating genetic parameters for the lamb survival trait (Van Wyk *et al.*, 2004; Everett-Hincks *et al.*, 2005, 2014) and for analyses involving other traits where the maternal genetic variance is estimated (Hagger, 1998; Al-Shorepy and Notter, 1998; Van Wyk *et al.*, 2004).

The ratio of litter variance to total phenotypic variance was greater than the direct and maternal genetic ratios in models B and LogB, indicating a common environmental or non-genetic resemblance among littermates. This is similar to what was reported in previous studies on lamb survival (Amer and Jopson, 2003; Van Wyk *et al.*, 2004; Everett-Hincks *et al.*, 2014).

Table 11. Estimates of variance components and genetic parameters (with their SE in parentheses), proportion, and restricted log-likelihood for the new survival trait (NEWSUR) using different linear models

	Model[b]			
Item[a]	A1	A2	A3	B
$\hat{\sigma}^2(BDAYc)$			0.007 (0.002)	0.007 (0.003)
$\hat{\sigma}^2(litter)$				0.014 (0.002)
$\hat{\sigma}^2(e)$	0.109 (0.002)	0.108 (0.002)	0.108 (0.002)	0.098 (0.002)
$\hat{\sigma}^2(a)$	0.007 (0.003)	0.006 (0.002)	0.005 (0.002)	0.005 (0.003)
$\hat{\sigma}(a,m)$	-0.001 (0.003)	-0.001 (0.002)	-0.001 (0.002)	-0.002 (0.002)
$\hat{\sigma}^2(m)$	0.008 (0.003)	0.008 (0.003)	0.007 (0.002)	0.002 (0.002)
$\hat{\sigma}^2(y)$	0.121 (0.003)	0.119 (0.003)	0.126 (0.004)	0.123 (0.004)
$\hat{h}^2(a)$	0.055 (0.000)	0.052 (0.000)	0.043 (0.000)	0.040 (0.002)
$\hat{h}^2(m)$	0.064 (0.000)	0.064 (0.000)	0.055 (0.000)	0.019 (0.001)
$\hat{\rho}(a,m)$	-0.186 (0.003)	-0.212 (0.003)	-0.136 (0.003)	-0.438 (0.003)
$\hat{\sigma}^2(g)$	0.012 (0.003)	0.011 (0.003)	0.011 (0.003)	0.004 (0.002)
$\hat{h}^2(g)$	0.097 (0.000)	0.092 (0.000)	0.085 (0.000)	0.035 (0.000)
LnL	-1564964.02	-1562810.38	-1558802.22	-1555568.89

[a] $\hat{\sigma}^2$(BDAYc) = estimated variance of the day of birth (classified in periods of 2 wk) × flock × year of birth effect; $\hat{\sigma}^2$(litter) = estimated variance of the litter effect; $\hat{\sigma}^2(e)$ = estimated residual variance; $\hat{\sigma}^2(a)$ = estimated direct additive genetic variance; $\hat{\sigma}(a,m)$ = estimated covariance between direct and maternal effects; $\hat{\sigma}^2(m)$ = estimated genetic maternal additive genetic variance; $\hat{\sigma}^2(y)$ = estimated phenotypic variance; $\hat{h}^2(a)$ = estimated direct heritability; $\hat{h}^2(m)$ = estimated maternal heritability; $\hat{\rho}(a,m)$ = estimated genetic correlation between direct and maternal effects; $\hat{\sigma}^2(g)$ = estimated total genetic (direct + maternal) variance; $\hat{h}^2(g)$ = estimated total heritability, LnL = restricted log-likelihood.

[b] A1 = model with no day of birth fitted; A2 = model where the interaction between effects day of birth, flock, and year of birth is fitted as fixed covariates; A3 = model where the interaction between day of birth classes, flock, and year of birth is fitted as random effects; B = same as A3 but with an extra random litter effect fitted.

Table 12. Estimates of variance components and genetic parameters (with their SE in parentheses), and restricted log-likelihood for the new survival trait (NEWSUR) using 2 different logistic models (LogA and LogB)

	Model[b]	
Item[a]	LogA	LogB
$\hat{\sigma}^2(BDAYc)$	0.358 (0.015)	0.347 (0.015)
$\hat{\sigma}^2(litter)$		0.355 (0.010)
$\hat{\sigma}^2(e)$	3.290 (0.000)	3.290 (0.000)
$\hat{\sigma}^2(a)$	0.128 (0.007)	0.083 (0.006)
$\hat{\sigma}(a,m)$	0.006 (0.010)	0.002 (0.007)
$\hat{\sigma}^2(m)$	0.315 (0.012)	0.141 (0.010)
$\hat{\sigma}^2(y)$	4.102 (0.019)	4.220 (0.019)
$\hat{h}^2(a)$	0.031 (0.002)	0.020 (0.002)
$\hat{h}^2(m)$	0.077 (0.003)	0.033 (0.002)
$\hat{\rho}(a,m)$	0.028 (0.052)	0.023 (0.068)
$\hat{\sigma}^2(g)$	0.454 (0.013)	0.229 (0.011)
$\hat{h}^2(g)$	0.111 (0.003)	0.054 (0.003)
LnL	-1454703.73	-1402085.36

[a] $\hat{\sigma}^2$(BDAYc) = estimated variance of the day of birth (classified in periods of 2 wk) × flock × b effect; $\hat{\sigma}^2$(litter) = estimated variance of the litter effect; $\hat{\sigma}^2(e)$ = estimated residual variance; $\hat{\sigma}^2(a)$ = estimated direct additive genetic variance; $\hat{\sigma}(a,m)$ = estimated covariance between direct and maternal effects; $\hat{\sigma}^2(m)$ = estimated genetic maternal additive genetic variance; $\hat{\sigma}^2(y)$ = estimated phenotypic variance; $\hat{h}^2(a)$ = estimated direct heritability; $\hat{h}^2(m)$ = estimated maternal heritability; $\hat{\rho}(a,m)$ = estimated genetic correlation between direct and maternal effects; $\hat{\sigma}^2(g)$ = estimated total genetic (direct + maternal) variance; $\hat{h}^2(g)$ = estimated total heritability; LnL = restricted log-likelihood.

[b] LogA = model where the interaction between day of birth classes, flock, and year of birth is fitted as random effects and where phenotypes are transformed using a logit transformation; LogB = same as model LogA but with an extra random litter effect fitted.

The h^2 attributed to direct effects ranged from 2 to 5.5% for all models. Model A1 generated the greatest direct h^2 whereas the LogB model generated the lowest h^2 when the same data set was used. The h^2 attributed to maternal effects ranged from 1.9 to 7.7% for all models. The greatest maternal h^2 was reported with model LogA and was lowest for model B. These values were low but in agreement with literature. A study involving 7 commercial New Zealand sheep flocks provided h^2 estimates for lamb survival ranged from 0.8 to 7% for direct effects and from 0.2 to 7.5% for maternal effects (Amer and Jopson, 2003). Safari *et al.* (2005) reviewed h^2 estimates for lamb survival trait from 16 studies and reported an averaged h^2 of 3 and of 5% for direct and maternal effects, respectively. Hatcher *et al.* (2010) found heritabilities ranged from 2 to 5% for direct effects and from 3 to 7% for maternal effects in Australian Merino sheep.

The maternal and direct genetic effects appeared to be negatively and slightly to moderately correlated (from -0.14 to -0.44) for models without phenotype transformation. This negative genetic correlation suggests that some genes that support survival in the individual may also contribute to poor survival in their progeny and vice-versa. However, Robinson, (1996) showed, using simulated data, that negative estimates of correlation between the direct and the maternal genetic effects may be obtained in the absence of a true antagonism between them. For both logistic models, the genetic correlation between direct and maternal effects was positive but low (from 0.02 to 0.03). The sign of the correlation differed between the logistic models and the linear models but the estimates of correlation had a large SE in logistic models. Therefore, the correlation between direct and maternal genetic effects for NEWSUR was significant for linear models and not significant for logistic models. Some previous studies on lamb survival (Morris *et al.*, 2000; Everett-Hincks *et al.*, 2005, 2014; Cloete *et al.*, 2009) also reported negative correlations between direct and maternal genetic effects using linear or logistic models. In a study of lamb survival in Rambouillet and Finnsheep with linear and threshold models, Matos *et al.* (2000) obtained positive correlations, albeit with large SE too.

COMPARISON OF MODELS BASED ON THEIR PREDICTIVE ABILITY

Model selection of random effects was assessed for inclusion based on log-likelihood values. Log-likelihood values obtained for NEWSUR with no transformation of phenotypes are reported in Table 11, and those for logistic models are reported in Table 12. With or without a logit transformation of NEWSUR, the models that included litter effect (*i.e.* models B and LogB) were superior to the models without this effect. Inclusion of the random effect of litter (dam/year) in survival analyses in populations where multiple births are prevalent is relatively common (Van Arendonk *et al.*, 1996; Knol *et al.*, 2002). Therefore, the inclusion of litter effects in the genetic evaluation model for NEWSUR is required to avoid inappropriate weighting of the contribution of lamb survival from multiple births when calculating EBV of NEWSUR.

Log-likelihood values cannot be used to compare the linear and logistic models; therefore, MSE were calculated, as described in the methodology, to assess whether model LogB was a better fit for lamb survival trait when compared with the model B. The MSE were slightly lower with the linear model B (from 0.1275 to 0.1284) than the logistic model LogB (from 0.1283 to 0.1293) and much lower than the intercept model (ranged from 0.1624 to 0.1672). Accordingly, the linear model B seems to be a better fit for NEWSUR than the model with a logit transformation (model LogB).

CONCLUSIONS

According to the results of this large study, the new decision rules to assign NEWSUR could reclaim and include more records in the genetic analysis of LS. Variance or covariance components and derived genetic parameters for NEWSUR were estimated with linear and logistic models. The correlation between direct and maternal genetic effects was negative for the linear models and positive but not significantly different from zero for the logistic models. The h^2 estimates were low but consistent with those found in other studies on lamb survival in sheep. This study also showed that the maternal heritabilities may be overestimated when the litter random effect was not fitted. Among the tested models, models fitting the random effect of litter effect displayed the best fit for NEWSUR according to their log-likelihoods values (*i.e.* models B and LogB). In terms of predictive ability, model B performed slightly better than model LogB. Thus, model B, which is the linear model fitting the random litter effect, would be the model of choice to implement the genetic evaluation of NEWSUR for the New Zealand sheep to improve the genetic progress for lamb survival to weaning.

LITERATURE CITED

Al-Shorepy, S. A., and D. R. Notter. 1998. Genetic parameters for lamb birth weight in spring and autumn lambing. Anim. Sci. 67:327–332.

Amer, P. R., and N. B. Jopson. 2003. Genetics of lamb survival. Proc. New Zealand Society of Animal Production. Vol. 63. p. 201–203.

Brien, F. D., S. W. P. Cloete, N. M. Fogarty, J. C. Greeff, M. L. Hebart, S. Hiendleder, J. E. H. Edwards, J. M. Kelly, K. L. Kind, D. O. Kleemann, K. L. Plush, and D. R. Miller. 2014. A review of the genetic and epigenetic factors affecting lamb survival. Anim. Prod. Sci. 54:667-693.

Cloete, S. W. P., I. Misztal, and J. J. Olivier. 2009. Genetic parameters and trends for lamb survival and birth weight in a Merino flock divergently selected for multiple rearing ability. J. Anim. Sci. 87:2196–2208.

Everett-Hincks, J. M., and K. G. Dodds. 2008. Management of maternal-offspring behaviour to improve lamb survival in easy care sheep systems. J. Anim. Sci. 86:E259–E270.

Everett-Hincks, J. M., N. Lopez-Villalobos, H. T. Blair, and K. J. Stafford. 2005. The effect of ewe maternal behaviour score on lamb and litter survival. Livest. Prod. Sci. 93:51–61.

Everett-Hincks, J. M., H. C. Mathias-Davis, G. J. Greer, B. Auvray, and K. G. Dodds. 2014. Genetic parameters for lamb birth weight, survival and death risk traits. J. Anim. Sci. 92:2885–2895.

Gilmour, A. R., B. J. Gogel, B. R. Cullis, and R. Thompson. 2009. ASReml user guide release 3.0. VSN Int. Ltd Hemel Hempstead UK. Available from: http://vsni.de/downloads/asreml/release3/UserGuide.pdf

Hagger, C. 1998. Litter, permanent environmental, ram-flock, and genetic effects on early weight gain of lambs. J. Anim. Sci. 76:452–457.

Hatcher, S., K. D. Atkins, and E. Safari. 2010. Lamb survival in Australian Merino Sheep: A genetic analysis. J. Anim. Sci. 88:3198–3205.

Hawkins, B. W., and S. Wu. 2011. Understanding the variability in sheep and beef farm profitability. Proc. N. Z. Grassl. Assoc. 73:23–26.

Knol, E. F., J. I. Leenhouwers, and T. van der Lende. 2002. Genetic aspects of piglet survival. Livest. Prod. Sci. 78:47–55.

Lopez-Villalobos, N, and D.J. Garrick. 1999. Genetic parameter estimates for lamb survival in Romney sheep. Proc. New Zealand Society of Animal Production. Vol. 99. p. 121–124.

Matos, C. A. P., D. L. Thomas, L. D. Young, and D. Gianola. 2000. Genetic analyses of lamb survival in Rambouillet and Finnsheep flocks by linear and threshold models. J. Anim. Sci. 71:227–234.

McDermott, A., C. Saunders, E. Zellman, T. Hope, and A. Fisher. 2008. The Key Elements of Success and Failure in the NZ Sheep Meat Industry from 1980 - 2007. Agribusiness and Economics Research Unit, Lincoln University, Lincoln, N.Z. Available from: http://www.aren.org.nz/publications.asp

Morris, C. A., S. M. Hickey, and J. N. Clarke. 2000. Genetic and environmental factors affecting lamb survival at birth and through to weaning. N. Z. J. Agric. Res. 43:515–524.

Price, S. 2000. What are the crossbreeding options? In: Crossbreeding for New Zealand sheep farmers. CF2000. Sheep Council, Lawrence Otago, New Zealand. p. 19–23.

Robinson, D. L. 1996. Models which might explain negative correlations between direct and maternal genetic effects. Livest. Prod. Sci. 45:111–122.

Safari, E., N. M. Fogarty, and A. R. Gilmour. 2005. A review of genetic parameter estimates for wool, growth, meat and reproduction traits in sheep. Livest. Prod. Sci. 92:271–289.

Van Arendonk, J. A. M., C. van Rosmeulen, L. L. G. Janss, and E. F. Knol. 1996. Estimation of direct and maternal genetic (co) variances for survival within litters of piglets. Livest. Prod. Sci. 46:163–171.

Van Wyk, J. B., M. D. Fair, and S. W. P. Cloete. 2004. Revised models and genetic parameter estimates for production and reproduction traits in the Elsenburg Dormer sheep stud. South Afr. J. Anim. Sci. 33:213–222.

CHAPTER IV. DEVELOPMENT OF GENETIC EVALUATION SYSTEMS IN A MULTI-BREED (CROSSBRED) POPULATION

CHAPTER IV. DEVELOPMENT OF GENETIC EVALUATION SYSTEMS IN A MULTI-BREED (CROSSBRED) POPULATION

OUTLINE

As mentioned in Chapter II, crossbreeding is a common practice in beef, poultry and swine productions. With the exception of New Zealand, crossbreeding of dairy cattle has garnered limited acceptance worldwide. For several years, however, crossbreeding of dairy cattle has become a topic of some considerable interest in response to concerns by dairy breeders and producers about declining fertility, health and survival in most major dairy breeds. By crossbreeding, dairy farmers aim to capitalise on breed complementarity and heterosis as well as remove the negative effects associated with the increasing levels of inbreeding within those breeds. However, most of genetic evaluation systems in dairy cattle compare animals only within breed or disregarding breed (composition) differences, as is the case in the systems developed and discussed in Chapter III. As a second step in the research strategy of this thesis, this chapter focused on the development of genetic evaluation models adapted to populations with various breed compositions in order to provide across-breed selection and mating programs to dairy breeders and producers. Firstly, a multi-breed model was developed to analyse milk yields from New Zealand dairy cattle, which is composed of a large proportion of crossbred cows (Chapter IV.1). Then, this multi-breed model was adapted and assessed for the joint genetic analysis of calving ease for purebred and crossbred animals from Holstein and Belgian Blue Walloon cattle (Chapter IV.2).

CHAPTER IV.1. ESTIMATION OF TEST-DAY MODEL (CO)VARIANCE COMPONENTS ACROSS BREEDS USING NEW ZEALAND DAIRY CATTLE DATA

Based on: Vanderick, S., B. L. Harris, J. E. Pryce, and N. Gengler. 2009. Estimation of test-day model (co)variance components across breeds using New Zealand dairy cattle data. J. Dairy Sci. 92:1240-1252.

CHAPTER IV.1. ESTIMATION OF TEST-DAY MODEL (CO)VARIANCE COMPONENTS ACROSS BREEDS USING NEW ZEALAND DAIRY CATTLE DATA

OUTLINE

Chapter II stated that multi-breed models allow purebred animals to be evaluated more accurately for performance in crossbreeding compared to strictly single-breed models because they use all available progeny information (*i.e.* crossbred and purebred data). As a result, a greater response to selection for crossbred performance can occur when a combined crossbred-purebred genetic evaluation is performed. To do this, genetic correlations of additive effects in different breeds need to be known. However, literature results for genetic correlations among additive effects from different dairy breeds are scarce. Therefore, the objective of this chapter was to develop a multi-breed model adapted to a multi-breed dairy cattle population to model different additive effects according to breed composition in order to estimate genetic correlations across breeds. This study was conducted on first-lactation milk yield records from New Zealand dairy cattle, in which a large proportion of animals are crossbred, mostly Holstein-Friesians × Jersey. These data were provided by the Livestock Improvement Corporation (LIC), Hamilton, New Zealand. The originality of this study lies in the use of a random regression test-day model, in which random regressions were breed-specific for each random effect. First, (co)variance components were estimated for purebred and crossbred animals using the test-day model. Then, derived genetic parameters within and across breed were explored.

ABSTRACT

In New Zealand, a large proportion of cows are currently crossbreds, mostly Holstein-Friesians (HF) × Jersey (JE). The genetic evaluation system for milk yield is considering the same additive genetic effects for all breeds. The objective was to model different additive effects according to parental breeds to obtain first estimates of correlations among breed-specific effects and to study the usefulness of this type of random regression test-day model. Estimates of (co)variance components for purebred HF and JE cattle in purebred herds were computed by using a single-breed model. This analysis showed differences between the 2 breeds, with a greater variability in the HF breed. (Co)variance components for purebred HF and JE and crossbred HF × JE cattle were then estimated by using a complete multi-breed model in which computations of complete across-breed (co)variances were simplified by correlating only eigenvectors for HF and JE random regressions of the same order as obtained from single-breed analysis. Parameter estimates differed more strongly than expected between the single-breed and multi-breed analyses especially for JE. This could be due to differences between animals and management in purebred and nonpurebred herds. In addition, the model used only partially accounted for heterosis The multi-breed analysis showed additive genetic differences between the HF and JE breeds, expressed as genetic correlations of additive effects in both breeds, especially in linear and quadratic Legendre polynomials (respectively, 0.807 and 0.604). The differences were small for overall milk production (0.926). Results showed that permanent environmental lactation curves were highly correlated across breeds; however intraherd lactation curves were also affected by the breed-environment interaction. This result may indicate the existence of breed-specific competition effects that vary through the different lactation stages. In conclusion, a multi-breed model similar to the one presented could optimally use the environmental and genetic parameters and provide breed-dependent additive breeding values. This model could also be a useful tool to evaluate crossbred dairy cattle populations like those in New Zealand. However, a routine evaluation would still require the development of an improved methodology. It would also be computationally very challenging because of the simultaneous presence of a large number of breeds.

Keywords: crossbreeding, Holstein-Friesian, Jersey, multi-breed model

INTRODUCTION

Crossbreeding is a method used for improving animal production in pigs, beef cattle, and poultry (*e.g.* Wei and van der Werf, 1995). However, it has not been widely used in dairy cattle in most temperate countries until recently because of the high milk production of the Holstein-Friesian (HF) breed (Touchberry, 1992). Most current purebred HF populations in the world were created by upgrading existing European Friesian or similar populations. Harris and Kolver (2001) gave a more detailed history of the New Zealand HF (NZHF) population that showed the different phases. The original NZHF population was first developed from animals imported from the West Coast of the United States before 1925, and it then remained as a closed population. Most of the cattle before 1960 (approximately 75%) were Jersey (JE). Therefore, until the early 1980s upgrading was first from JE to NZHF by using locally available sires. The descendants of this process remained lighter than overseas HF (OSHF) from North America or Europe. Since the 1980s, OSHF sires have become more popular. However, the daughters of these animals are heavier and seem to be less fertile and have decreased survival rates (Harris and Kolver, 2001). Therefore, most New Zealand dairy farmers have not upgraded to OSHF, and as a result of this experience, there has been even more widespread popularity of crossbreeding. Even though dairy farmers in North America and Europe were very reluctant to crossbreed at that time, crossbreeding has been a feature of the history of the dairy industry of New Zealand. Currently, more than one-third of dairy replacements are crossbred, mostly HF × JE. This is because both breeds, the HF and JE, are economically comparable under typical New Zealand production systems (*e.g.* Lopez-Villalobos and Garrick, 2002), and complementarities characteristics from these breeds and heterosis effects favour crossbreds (*e.g.* Lopez-Villalobos and Garrick, 2002). Crossbreeding provides a good opportunity to maximize the net income per hectare under New Zealand pastoral conditions by improving fertility and survival, and also by improving (or altering) the composition of milk, which is very important because the dairy company payment structures reward farmers for the amount of milk solids (fat and protein) produced. Therefore, crossbreds benefit from the high volumes achieved in the HF breed as well as from the beneficial fat and protein composition of the JE breed (Montgomerie, 2002). Several studies have demonstrated the economic superiority of crossbred HF × JE cows in New-Zealand (*e.g.* Lopez-Villalobos and Garrick, 1996, 1997; Lopez-Villalobos *et al.*, 2002) and potentially elsewhere (*e.g.* VanRaden and Sanders, 2001).

Additionally, crossbred bulls are currently being progeny tested by the major New Zealand breeding companies in response to farmer requests. Farmers are willing to accept a reduction in heterosis in order to have a type of cow that they recognize as being the most profitable for their system, even when heterosis effects are ignored.

From 1996 up to the start of 2007, New Zealand dairy cattle were genetically evaluated across breeds for yield traits using a 2-step test-day model, in which test-day production records were combined to predict 270-day yields, and an animal model was used for the genetic evaluation of these predicted yields (Harris, 1994, 1995). A new test-day model has been developed to use the milk production herd-test data to estimate breeding values, and this new system has been implemented in February 2007 (Harris *et al.*, 2006). The evaluated production traits are then included in an economic index called Breeding Worth (BW), which describes animal profitability per unit of feed (this is a feature of the across-breed evaluation) (Garrick *et al.*, 1997; Johnson, 1996). Heterosis is modeled as a fixed effect in the model correcting for mean differences. Thus, genetic contributions from purebreds to crossbreds are only partly taken into account; therefore, it does not allow an optimal use of crossbred data. Moreover, a study by Wei and Van der Werf (1995) showed that an optimal use of crossbred information jointly with purebred information in selection could bring more genetic progress in crossbreds. To do this, genetic correlation of additive effects in different breeds has to be known. Such results are rare, if not non-existent, in dairy cattle, whereas in swine or in beef cattle such results are less unusual (*e.g.* Lutaaya *et al.*, 2001; Newman *et al.*, 2002; Roso *et al.*, 2005; Zumbach *et al.*, 2007).

The objective of this study was to model different additive effects according to breed composition in order to estimate correlations among breed-specific effects. The second objective was also to study the usefulness of this type of a more complicated random regression test-day model.

MATERIALS AND METHODS

DATA

Data were provided by the Livestock Improvement Corporation (LIC), Hamilton, New Zealand, and comprised of records on cows from dedicated progeny-testing herds. The original data comprised 223 141 animals in production and a total of 500 134 animals in the pedigree.

Only animals of HF and JE inheritance were kept to estimate (co)variance components; therefore, animals whose summed proportion of HF and JE genes was lower than 100% were eliminated. No distinction was made among HF animals with different compositions of OSHF and NZHF genes; they were considered to belong to the same breed. This was also done to limit the study to a 2-breed situation. Data were limited to first-lactation test-day milk yield, which were recorded for 208 164 cows in 3481 herds with test days, equally spaced across the lactation, at 2 months intervals. After these edits, the breed composition of animals with records in first lactation were approximately 54% HF, 21% JE, and 25% HF × JE, and the distribution of herds per breed composition showed that 65% of herds had HF × JE, HF, and JE or were entirely composed of HF × JE animals, 25 % of herds were purebred HF and only 10% were purebred JE. In the context of this article, purebred means at least 95% of HF or JE genes. Crossbreds consequently were animal with a major breed composition of between 50% (included) and 95% (excluded) HF or JE genes.

Data sets were constructed based on a stratification of herds as a function of their average breed proportions. Herds with average breed compositions of 95% and more HF or JE genes, were considered purebred herds. Similarly, herds with an average breed composition of between 50% (included) and 95% (excluded) HF or JE genes were considered crossbred herds. Within the purebred herds, only animals whose proportions of HF or JE genes were 95% or more were selected. Because the data set was still too large for analysis, additional samplings were made on a herd base, in which animals within herds were kept together and randomly assigned to data sets. For purebred herds, 3 HF and 2 JE data sets were created. For crossbred herds, 5 data sets were created, composed of purebred and crossbred animals. Table 13 lists the number of herds, animals in production, test-day records, and records per lactation for the purebred and crossbred data sets used in this study. Means and standard deviation of daily milk yield are also given.

Table 13. Composition, daily milk yield average, and standard deviations of purebred and crossbred samples[1] used in this study

			Test-day records		Daily milk yield, kg	
Data sets[a]	Herds, n	Cows in production, n	n	Average per lactation, n	Average	SD
HF 1	160	11 721	44 531	3.8	13.30	4.35
HF 2	166	11 952	45 326	3.8	13.08	4.38
HF 3	165	11 748	45 324	3.9	13.68	4.46
JE 1	107	7703	29 512	3.8	9.94	3.14
JE 2	105	7671	30 605	4.0	9.76	3.03
HF×JE 1	64	4966	19 048	3.8	12.78	4.22
HF×JE 2	48	4918	19 769	4.0	12.50	4.20
HF×JE 3	44	4990	19 486	3.9	12.85	4.33
HF×JE 4	60	4959	18 305	3.7	12.62	4.21
HF×JE 5	64	4968	18 017	3.6	12.11	4.20

[a] HF n = Holstein-Friesian purebred sample n; JE n = Jersey purebred sample n; HF × JE n = Holstein-Friesian × Jersey crossbred sample n.

MODELS

The research was conducted in 2 steps. First, estimation of (co)variance components for purebred HF and JE cattle was done by using a simple single-breed model. Hereafter, this study is called hereafter a single-breed analysis. (Co)variance components for purebred HF and JE and crossbred HF × JE cattle were then estimated using a complete multi-breed model. This study was considered to be an across-breed analysis, hereafter called a multi-breed analysis. The reasons for using this strategy were 2-fold. First, we wanted to simplify actual estimations by estimating within and then across breeds. Second, this strategy allowed us to compare the results from both analyses. We acknowledge that there were 2 types of herds (purebred *vs* crossbred), implying that differences should be expected.

Single-breed analysis

(Co)Variance components were estimated for HF and JE breeds separately by using the purebred data sets. The single-breed model used was a random regression test-day model, written as:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Q}_h\mathbf{h} + \mathbf{Q}_a\mathbf{a} + \mathbf{Q}_p\mathbf{p} + \mathbf{e} \qquad (46)$$

where $\mathbf{y}$ is a vector of first-lactation test-day yields, $\mathbf{b}$ a vector of fixed effects for herd × test-day date, lactation stage (defined as classes of 5 DIM), gestation stage (definition based on 24 d carried calf classes), and regressions on age at calving within lactation periods (defined as 4 classes based on DIM : 5 to 64, 65 to 119, 120 to 184, and greater than 184) × calving season (March to June, July, August, September to December) × calving period (before 1994, 1994 to 1998, and after 1998) classes; and $\mathbf{h}$, $\mathbf{a}$ and $\mathbf{p}$ are vectors of unknown herd × calving year, permanent environmental, and additive genetic random regression effects; $\mathbf{Q}_h, \mathbf{Q}_a$ and $\mathbf{Q}_p$ are the covariate matrices for the third-order Legendre polynomials; linking $\mathbf{h}$, $\mathbf{a}$, and $\mathbf{p}$ with $\mathbf{y}$ and $\mathbf{e}$ is a vector of unknown residuals. The (co)variance matrices associated with the 3 random effects were $Var(\mathbf{h}) = \mathbf{I}_h \otimes \mathbf{H}$, $Var(\mathbf{p}) = \mathbf{I}_p \otimes \mathbf{P}$, and $Var(\mathbf{a}) = \mathbf{A}_a \otimes \mathbf{G}$ where $\mathbf{H}, \mathbf{P}, \text{and}\, \mathbf{G}$ were the elementary (co)variance matrices across the 3 Legendre polynomials. A random herd × calving year period effect was included in order to model potential additional environmental covariances because of a common herd and time-dependent effect. This random effect was defined as a combination between herds and 7 periods of 2 years of calving from 1989 to 2002. Hereafter, this effect is called the herd-period. The (co)variances among residuals were modeled as $Var(\mathbf{e}) = \mathbf{I}_n\sigma_e^2$, where σ_e^2 was the residual variance and $\mathbf{I}_n$ was an identity matrix of dimension n, where n is the number of records. By keeping this variance constant, we modeled differences in environmental variances across DIM by the other environmental random effects.

Multi-breed analysis

The model used for the multi-breed analysis was similar to those used for the single-breed studies, with 2 differences. In the definition of fixed effects, regressions on age at calving were computed within breed × lactation period × calving season × calving period classes, where breed was defined as previously: purebred HF, purebred JE or crossbred HF × JE. Effects of heterosis and recombination should be partially taken into account by this effect. All other fixed effects were defined without considering the breed of animals. The second difference was in the definition of random effects. Based on the results from the within-breed analysis and to simplify the estimation of (co)variance components across breeds, initial Legendre polynomials were transformed into three new independent regressions contained in the covariate matrices $\mathbf{Q}_{Th}, \mathbf{Q}_{Ta}$, and $\mathbf{Q}_{Tp}$ by using diagonalizations based on the following decompositions: $\mathbf{H} = \mathbf{T}_h \mathbf{D}_h \mathbf{T}_h', \mathbf{P} = \mathbf{T}_p \mathbf{D}_p \mathbf{T}_p'$, and $\mathbf{G} = \mathbf{T}_a \mathbf{D}_a \mathbf{T}_a'$. Original covariance structures were transformed to become $Var(\mathbf{h}_T) = \mathbf{I}_h \otimes \mathbf{D}_h$, $Var(\mathbf{p}_T) = \mathbf{I}_p \otimes \mathbf{D}_p$ and $Var(\mathbf{a}_T) = \mathbf{A}_a \otimes \mathbf{D}_a$ based on $Var(\mathbf{h}) = \mathbf{I}_h \otimes \mathbf{T}_h \mathbf{D}_h \mathbf{T}_h'$, $Var(\mathbf{p}) = \mathbf{I}_p \otimes \mathbf{T}_p \mathbf{D}_p \mathbf{T}_p'$ and $Var(\mathbf{a}) = \mathbf{A}_a \otimes \mathbf{T}_a \mathbf{D}_a \mathbf{T}_a'$.

To achieve equivalence between the original and the transformed model, transformed regression matrices required to model a record i were defined as $\mathbf{Q}_{Th(i)} = \mathbf{Q}_{h(i)} \mathbf{T}_h, \mathbf{Q}_{Ta(i)} = \mathbf{Q}_{a(i)} \mathbf{T}_a$, and $\mathbf{Q}_{Tp(i)} = \mathbf{Q}_{p(i)} \mathbf{T}_p$. Equivalence can be shown by rewriting the original model (38) for a single observation i:

$$y_{(i)} = \mathbf{X}_{(i)} \mathbf{b}_{(i)} + \mathbf{Q}_{h(i)} \mathbf{h}_{(i)} + \mathbf{Q}_a \mathbf{a}_{(i)} + \mathbf{Q}_p \mathbf{p}_{(i)} + e_{(i)} \quad (47)$$

The transformed model becomes

$$y_{(i)} = \mathbf{X}_{(i)} \mathbf{b}_{(i)} + \mathbf{Q}_{Th(i)} \mathbf{h}_{T(i)} + \mathbf{Q}_{Ta(i)} \mathbf{a}_{T(i)} + \mathbf{Q}_{Tp(i)} \mathbf{p}_{T(i)} + e_{(i)} \quad (48)$$

where, based on the equivalence between models (47) and (48), the following formulas can be established:

$$Var\left(\mathbf{Q}_{Th(i)}\mathbf{h}_{T(i)}\right) = \mathbf{Q}_{Th(i)}Var\left(\mathbf{h}_{T(i)}\right)\mathbf{Q}'_{Th(i)} = \mathbf{Q}_{Th(i)}\mathbf{D}_{h}\mathbf{Q}'_{Th(i)} = \mathbf{Q}_{h(i)}\mathbf{T}_{h}\mathbf{D}_{h}\mathbf{T}'_{h}\mathbf{Q}'_{h(i)} = \mathbf{Q}_{h(i)}\mathbf{H}\mathbf{Q}'_{h(i)},$$

$$Var\left(\mathbf{Q}_{Ta(i)}\mathbf{a}_{T(i)}\right) = \mathbf{Q}_{Ta(i)}Var\left(\mathbf{a}_{T(i)}\right)\mathbf{Q}'_{Ta(i)} = \mathbf{Q}_{Ta(i)}\mathbf{D}_{a}\mathbf{Q}'_{Ta(i)} = \mathbf{Q}_{a(i)}\mathbf{T}_{a}\mathbf{D}_{a}\mathbf{T}'_{a}\mathbf{Q}'_{a(i)} = \mathbf{Q}_{a(i)}\mathbf{G}\mathbf{Q}'_{a(i)}, \text{ and}$$

$$Var\left(\mathbf{Q}_{Tp(i)}\mathbf{p}_{T(i)}\right) = \mathbf{Q}_{Tp(i)}Var\left(\mathbf{p}_{T(i)}\right)\mathbf{Q}'_{Tp(i)} = \mathbf{Q}_{Tp(i)}\mathbf{D}_{p}\mathbf{Q}'_{Tp(i)} = \mathbf{Q}_{p(i)}\mathbf{T}_{a}\mathbf{D}_{a}\mathbf{T}'_{a}\mathbf{Q}'_{p(i)} = \mathbf{Q}_{p(i)}\mathbf{P}\mathbf{Q}'_{p(i)}.$$

By using the transformation, the global single-breed model could therefore be rewritten as

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Q}_{Th}\mathbf{h}_{T} + \mathbf{Q}_{Ta}\mathbf{a}_{T} + \mathbf{Q}_{Tp}\mathbf{p}_{T} + \mathbf{e} \qquad (49)$$

Based on model (47), an appropriate multi-breed model considering separate random effects per breed would be written as

$$\mathbf{y} = \mathbf{Xb} + \Phi^{(HF)}\left(\mathbf{Q}_{h}\mathbf{h}^{(HF)} + \mathbf{Q}_{a}\mathbf{a}^{(HF)} + \mathbf{Q}_{p}\mathbf{p}^{(HF)}\right) + \Phi^{(JE)}\left(\mathbf{Q}_{h}\mathbf{h}^{(JE)} + \mathbf{Q}_{a}\mathbf{a}^{(JE)} + \mathbf{Q}_{p}\mathbf{p}^{(JE)}\right) + \mathbf{e} \quad (50)$$

where $\Phi^{(HF)}$ and $\Phi^{(JE)}$ are diagonal incidence matrices containing the fraction of breed compositions linking observations to random effects. Notation of fractions of the breed composition of animals is written as $\varphi_{(i)}^{(HF)}$ and $\varphi_{i}^{(JE)}$ for every record i, with values being obviously identical for all records of the same animal. The feature of this model was that the definition of the covariate matrices included a multiplication with the fraction of HF, respectively, JE breed contributions. To simplify computations and based on model (48), the following transformed model was used:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Q}_{Th}^{(HF)}\mathbf{h}_{T}^{(HF)} + \mathbf{Q}_{a}^{(HF)}\mathbf{a}_{T}^{(HF)} + \mathbf{Q}_{p}^{(HF)}\mathbf{p}_{T}^{(HF)} + \mathbf{Q}_{Th}^{(JE)}\mathbf{h}_{T}^{(JE)} + \mathbf{Q}_{a}^{(JE)}\mathbf{a}_{T}^{(JE)} + \mathbf{Q}_{p}^{(JE)}\mathbf{p}_{T}^{(JE)} + \mathbf{e} \quad (51)$$

where breed-specific regressions for a given record i are $\mathbf{Q}_{Th(i)}^{(HF)} = \varphi_{(i)}^{(HF)}\mathbf{Q}_{Th(i)}, \mathbf{Q}_{Ta(i)}^{(HF)} = \varphi_{(i)}^{(HF)}\mathbf{Q}_{Ta(i)}$, and $\mathbf{Q}_{Tp(i)}^{(HF)} = \varphi_{(i)}^{(HF)}\mathbf{Q}_{Tp(i)}$ for HF and $\mathbf{Q}_{Th(i)}^{(JE)} = \varphi_{(i)}^{(JE)}\mathbf{Q}_{Th(i)}, \mathbf{Q}_{Ta(i)}^{(JE)} = \varphi_{(i)}^{(JE)}\mathbf{Q}_{Ta(i)}$, and $\mathbf{Q}_{Tp(i)}^{(JE)} = \varphi_{(i)}^{(JE)}\mathbf{Q}_{Tp(i)}$ for JE.

Computations of complete across-breed (co)variances were simplified by correlating only eigenvectors for HF and JE of the same order. The associated (co)variance matrices among these transformed random effects could then be shown grouped by type of effects:

$$Var\begin{pmatrix}\mathbf{h}_T^{(HF)}\\ \mathbf{h}_T^{(JE)}\end{pmatrix} = \mathbf{I}_h \otimes \begin{pmatrix}\mathbf{D}_h^{(HF)} & \mathbf{D}_h^{(HFxJE)}\\ \mathbf{D}_h^{(HFxJE)} & \mathbf{D}_h^{(JE)}\end{pmatrix}, \qquad Var\begin{pmatrix}\mathbf{p}_T^{(HF)}\\ \mathbf{p}_T^{(JE)}\end{pmatrix} = \mathbf{I}_p \otimes \begin{pmatrix}\mathbf{D}_p^{(HF)} & \mathbf{D}_p^{(HFxJE)}\\ \mathbf{D}_p^{(HFxJE)} & \mathbf{D}_p^{(JE)}\end{pmatrix}, \quad \text{and}$$

$$Var\begin{pmatrix}\mathbf{a}_T^{(HF)}\\ \mathbf{a}_T^{(JE)}\end{pmatrix} = \mathbf{A}_a \otimes \begin{pmatrix}\mathbf{D}_a^{(HF)} & \mathbf{D}_a^{(HFxJE)}\\ \mathbf{D}_a^{(HFxJE)} & \mathbf{D}_a^{(JE)}\end{pmatrix}.$$

The diagonal matrices $\mathbf{D}^{(HFxJE)}$ for every type of random effects contained the (co)variances across breeds among transformed regressions. Transformations were done in a way that forced the order of eigenvectors for HF and JE to remain identical (largest to smallest). This procedure is not a rank reduction; but makes the computations more robust because it concentrates (co)variances among breeds into 3 values instead of 6. Initial tests showed that stable convergence of variance component estimation procedures required the use of this procedure. In addition, interpretation of the (co)variances among breeds was simplified. The basic assumption was that the eigenvectors stayed similar if we compare the single-breed and multi-breed models.

Heterogeneity of residual variances was expected for animals with different breed compositions. Therefore, breed composition was used to calculate weights in order to standardize residual variances and to adjust for the expected heterogeneous residual variances. The weight for a record i of a given animal was defined as

$$weight_{(i)} = \left[\frac{HF_p \times \sqrt{\left(\sigma_e^2\right)^{(HF)}} + JE_p \times \sqrt{\left(\sigma_e^2\right)^{(JE)}}}{\varphi_{(i)}^{(HF)} \times \sqrt{\left(\sigma_e^2\right)^{(HF)}} + \varphi_{(i)}^{(JE)} \times \sqrt{\left(\sigma_e^2\right)^{(JE)}}}\right]^2, \quad (52)$$

where HF_p and JE_p are the average breed proportion for breed HF and JE of the population used for the estimations; $\left(\sigma_e^2\right)^{(HF)}$ and $\left(\sigma_e^2\right)^{(JE)}$ are the residual variances estimated with the single-breed model; and $\varphi_i^{(HF)}$ and $\varphi_i^{(JE)}$ are the breed proportions for breed HF and JE of a record i of a given animal in production. The (co)variances among unknown residuals were modeled as $Var(\mathbf{e}) = \mathbf{W}\sigma_e^2$ where σ_e^2 was the multi-breed residual variance, and the matrix diagonal matrix $\mathbf{W}$ contained the inverse of the weight for every record.

ESTIMATION OF (CO)VARIANCE COMPONENTS

(Co)variance components were estimated by using REML. The expectation-maximization algorithm (EM) with acceleration EM-REML was used. This algorithm estimates the parameters by indirect approximation of the first derivative of the likelihood function. Expectation-maximization-REML is very stable but the convergence rate is very slow (Misztal, 2002). Therefore, the average information REML algorithm (Jensen *et al.*, 1997) was also used. It uses approximate second derivatives and is computationally more demanding, although it requires less iteration to converge. However, convergence problems appeared when the (co)variance matrices were not positive definite (Meyer, 1997; Misztal *et al.*, 2000) and most of the computations needed to be done by EM-REML, which proved to be more stable.

RESULTS AND DISCUSSION

SINGLE-BREED ANALYSIS

The (co)variance components were estimated for each sample, and from these estimates, the means and standard deviations of the samples for each breed were calculated. Results for variances and (co)variance for constant, linear, and quadratic random regression coefficients are presented in Table 14. In general; estimates were quite consistent across the 3 HF and 2 JE samples, except the herd-period effect in HF, which showed a greater variability.

Variances of random effects, except for the residual variances, which were considered to be constant, varied with DIM because of the use of random regression models. Phenotypic variances for milk as a function of DIM for HF and JE breeds are shown in Figure 5 5. The HF breed showed a greater variability than the JE breed, as confirmed by the herd-period, permanent environment, and genetic variances patterns of the 2 breeds (Figure 6). This can be explained mainly by a scale effect due to the greater production of HF. It is also likely that HF showed a greater genetic variability because of the importation of HF genetic material from overseas (Canada, United States, or Europe), especially since the 1980s, creating 2 different HF strains (a NZHF strain and an imported OSHF strain). Harris and Kolver (2001) have reviewed the effect of Holsteinization on pastoral dairy farming and have discussed the differences between the 2 HF strains in New Zealand. In this study, we did not distinguish between the 2 strains because the genetic difference between the 2 HF strains can be considered less important than the one between HF and JE. Animals of non-native JE origin are also present in New Zealand pedigrees, although to a much lesser degree.

Table 14. Mean and standard deviation of variances and (co)variances (liters²) for constant, linear, and quadratic random regression coefficients from single-breed models

	HF[1] purebreds						JE[2] purebreds					
	Constant		Linear		Quadratic		Constant		Linear		Quadratic	
Item	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Residual												
Constant	17.771	1.211					8.674	0.537				
Herd period												
Constant	8.138	3.840					6.624	0.445				
Linear	-4.035	1.530	3.945	1.665			-3.834	0.527	2.430	0.417		
Quadratic	-0.366	0.880	-0.500	0.376	0.415	0.365	1.503	0.403	-0.920	0.201	0.457	0.274
Permanent environment												
Constant	17.538	1.072					8.434	0.760				
Linear	-7.943	0.338	4.172	0.641			-3.701	0.112	2.055	0.170		
Quadratic	1.674	0.257	-1.971	0.400	3.268	0.097	0.601	0.227	-0.972	0.115	1.610	0.204
Genetic												
Constant	12.355	0.839					7.455	0.713				
Linear	-2.414	0.271	1.810	0.219			-1.136	0.414	1.031	0.112		
Quadratic	-0.500	0.347	0.013	0.193	0.484	0.048	-0.177	0.198	0.218	0.072	0.330	0.001

[1] HF = Holstein-Friesian.

[2] JE = Jersey.

Herd-period variance trends are given in Figure 6. The trajectory of this variance was less pronounced in HF but was generally greater than for JE. The differences may be due to a difference in management between the HF and JE purebred herds in New Zealand. For example, the greater stocking rates of the JE breed could be an explanation for these differences in variability. Another assumption is that the geographical location of HF and JE herds could have influenced the results. Bryant *et al.* (2007a) showed that HF and JE have reacted differently to differences in climate or altitude. Incidentally, purebred JE herds are more common in the northern part of the North Island, whereas purebred HF herds are more common in the cooler parts of the South Island.

Lactation variances were computed over a 270-day standard lactation. Table 15 presents the variances and relative values obtained when compared with phenotypic variance. The values confirmed the previous results in Figure 6 showing greater relative herd period variances and slightly greater relative genetic variances for JE. Residual variance on a lactation basis was very low because environmental correlations across DIM are taken into account by other environmental effects, but residual correlations among DIM were assumed to be zero.

Table 15. Variances (liters^2 × 10 000) and relative variances to phenotypic computed for a lactation of 270 DIM in the single-breed models of purebred Holstein-Friesian (HF) and Jersey (JE) animals

	HF purebreds		JE purebreds	
Type of variance	Variance	Relative	Variance	Relative
Phenotypic	31.80	1.00	18.40	1.00
Permanent environment	14.50	0.46	7.00	0.38
Genetic	9.91	0.31	5.87	0.32
Herd period	7.31	0.23	5.55	0.30
Residual	0.05	0.00	0.02	0.00

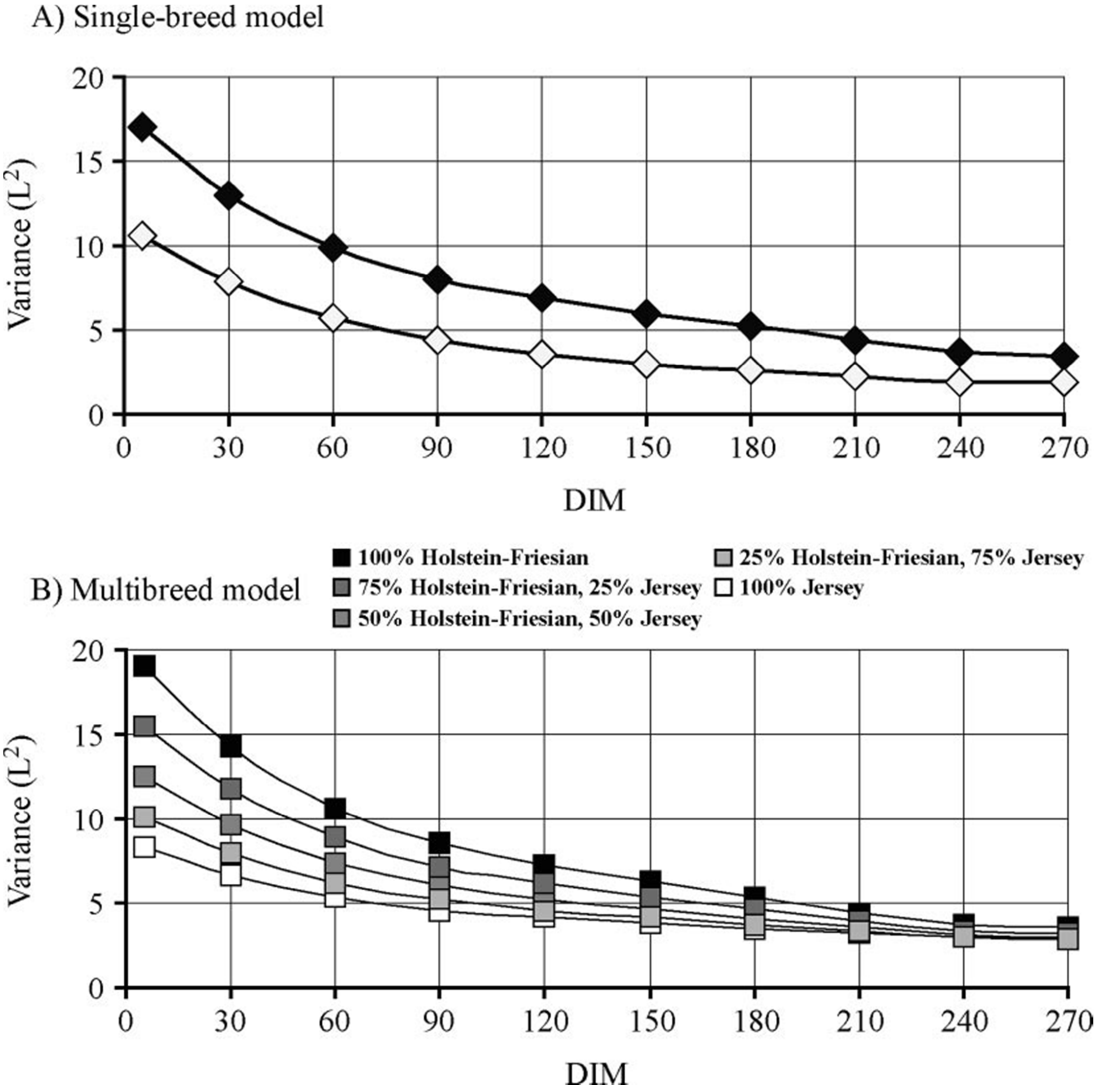
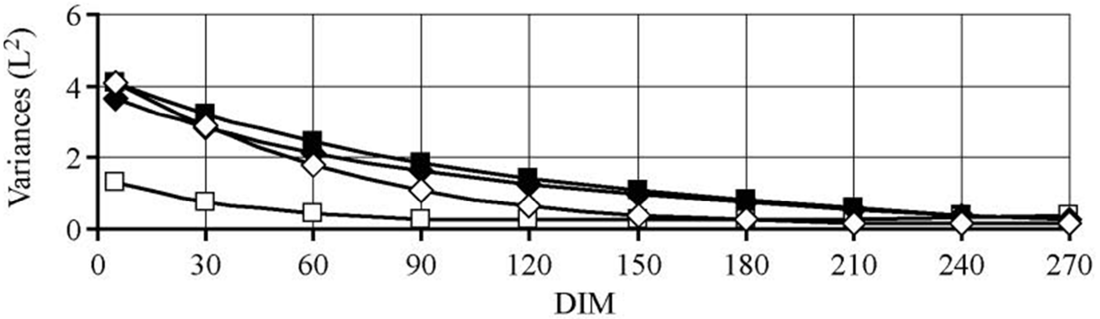
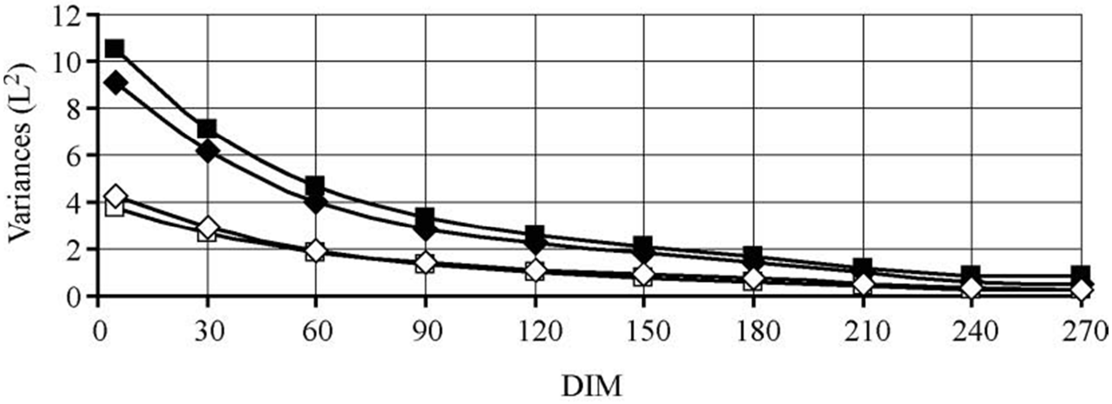


Figure 5. A) Comparison of phenotypic variances obtained by the single-breed analysis for Holstein-Friesian (♦) and Jersey (◊) animals for milk as function of DIM. B) Comparison of phenotypic variances obtained by the multi-breed analysis for purebred and crossbred animals.

A) Herd-period



B) Permanent environment



C) Additive genetic

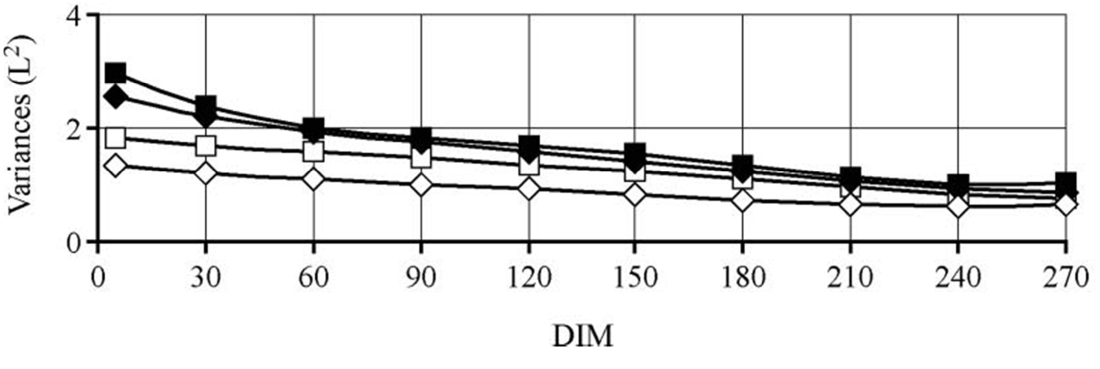


Figure 6. Herd-period (A), permanent environment (B), and additive genetic variances (C) for milk as function of DIM estimated by using the single-breed (◊, ♦) and multi-breed models (□, ■) for Holstein-Friesian (solid symbols) and Jersey (open symbols) animals

Table 16. Heritabilities, and genetic (above diagonal) and phenotypic (below diagonal) correlations for daily milk yield among first lactation, estimations from single-breed models for Holstein-Friesian and Jersey animals

	DIM				
Item	5	60	120	180	270
Holstein-Friesian					
5	0.150	0.954	0.844	0.739	0.575
60	0.804	0.197	0.965	0.896	0.670
120	0.597	0.719	0.231	0.977	0.747
180	0.415	0.581	0.672	0.241	0.841
270	0.361	0.401	0.408	0.458	0.255
Jersey					
5	0.126	0.960	0.880	0.816	0.622
60	0.840	0.192	0.976	0.925	0.630
120	0.648	0.746	0.259	0.978	0.665
180	0.471	0.608	0.684	0.285	0.782
270	0.466	0.479	0.457	0.447	0.350

Daily estimates of genetic and phenotypic correlations between different stages of lactation and daily estimates of heritabilities are presented in Table 16, with heritabilities on the diagonal, genetic correlations above the diagonal, and phenotypic correlations below the diagonal. As expected, these correlations became weaker as the interval between tests increased. On average, correlations were greater for JE than HF animals. Heritability (h^2) estimates increased from the beginning until the end of the lactation for both breeds. This trend was confirmed by the representation of heritabilities for milk yield as function of DIM, as illustrated in Figure 7. According to Figure 7, purebred HF animals had a greater h^2 than JE at the beginning of lactation, but from 60 DIM, the trend was reversed. As reported in Table 14, the h^2 for 270-day lactation yields (represented by the genetic relative variance) were similar for HF and JE purebreds (respectively, 0.31 and 0.32) and these values were slightly less than the h^2 estimate used for the genetic evaluation of milk in New Zealand (0.36; Interbull, 2007). These estimates were not totally in line with literature (*e.g.* Ahlborn and Dempfle, 1992; Interbull, 2007), in which values of heritabilities for the lactation milk yield trait for the JE breed tended to be greater than those for HF. For example, in the United States, heritabilities of 0.25 to 0.35 are used for the Holstein breed and 0.30 to 0.40 for the JE breed (Interbull, 2007). However, there may be other reasons to explain our results.

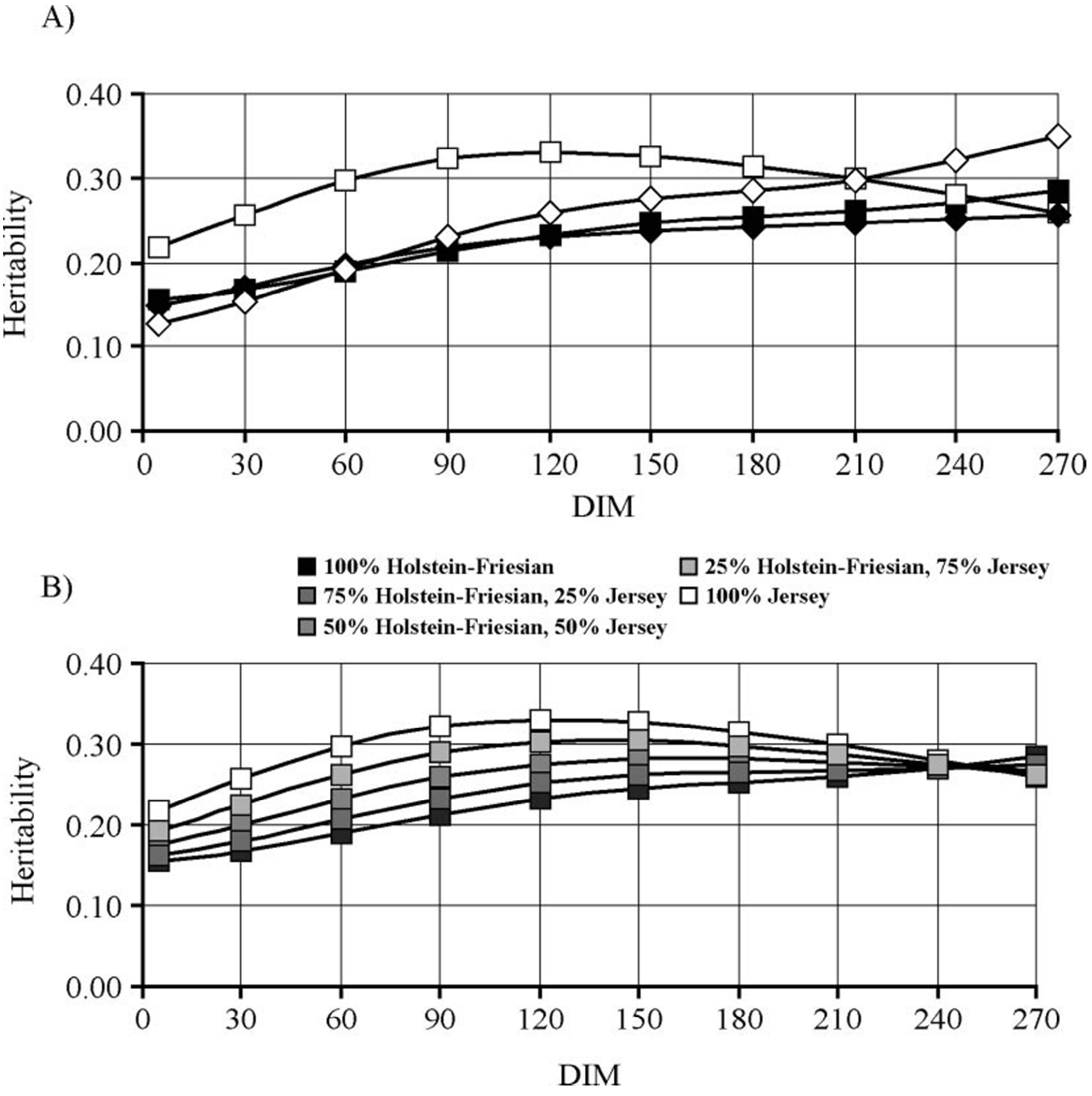


Figure 7. Heritabilities for milk as a function of DIM estimated by using single-breed (◊, ♦) and multi-breed models (□, ■) for purebred Holstein-Friesian (solid symbols) and Jersey (open symbols) animals (A) and for purebred and crossbred animals using the multi-breed model (B).

MULTI-BREED ANALYSIS

As for the single-breed analysis, (co)variance components estimated for each sample were averaged. Means and standard deviations of (co)variances for constant, linear and quadratic random regression coefficients are given in Table 17. The patterns and magnitude of h^2 estimates were different between the single-breed and multi-breed analyses, especially for JE purebred animals on a daily basis (see Figure 7). Values of variances and h^2 estimates computed for 270-day lactation are presented in Table 18. These results were also quite different (mainly for JE animals) from those obtained in the single-breed models. They were more in line with literature, albeit slightly lower (*e.g.* Ahlborn and Dempfle, 1992; Interbull, 2007). As confirmed by the values in Table 15 and Table 18, the lactation h^2 for JE changed from 0.32 to 0.52 when changing from the single-breed to the multi-breed model. Figure 6 shows that a greater genetic variance and a lower herd-period estimate from the multi-breed analysis for JE animals induced the difference in h^2. Interpretation of these large differences is not easy. A first potential reason could be the difference in the types of animals in each study because completely different data sets were used. In addition, the management between purebred herds and herds considered crossbred could partially explain these differences. This hypothesis is supported by the change in the JE herd-period variance with a decrease in the estimated value from 5.55 (single-breed) to 1.45 (multi-breed). The JE herd-period variance was, in fact, as much as 4 times greater, mainly for the first half of the lactation, in the single-breed model than in the multi-breed model. Few differences in permanent environmental variances were observed between the single-breed model and the multi-breed model (see Figure 6). Finally, the multi-breed model used did not fully account for heterosis. This could also have inflated observed genetic variance. This hypothesis was supported by the observed increase in genetic variance for JE when comparing the single-breed and multi-breed models, where 5.87 became 8.85.

Table 17. Mean and standard deviation of variances and (co)variances (liters²) for constant linear and quadratic random regression coefficients from the multi-breed model

	HF[1] crossbreds						JE[2] crossbreds					
	Constant		Linear		Quadratic		Constant		Linear		Quadratic	
Item	Estimate	SD	Estimate	SD	Estimate	SD	Estimate	SD	Estimate	SD	Estimate	SD
Residual												
Constant	16.738	1.346					11.694	0.940				
Herd period												
Constant	9.331	4.193										
Linear	-4.842	2.682	4.299	1.630								
Quadratic	-0.357	0.250	-0.468	0.420	0.334	0.219						
Constant	3.172	1.629	-1.396	1.081	-0.228	0.319	1.890	0.958				
Linear	-1.424	1.025	1.590	1.075	-0.349	0.436	-0.159	1.180	1.548	1.063		
Quadratic	0.497	0.417	-0.725	0.442	-0.020	0.258	0.035	0.508	-0.449	0.524	0.702	0.270
Permanent environment												
Constant	20.289	1.205										
Linear	-8.876	0.527	5.415	0.445								
Quadratic	1.998	0.197	-2.055	0.116	3.960	0.244						
Constant	12.617	1.231	-5.663	0.562	1.371	0.087	8.016	1.477				
Linear	-5.552	0.548	3.169	0.286	-1.280	0.237	-3.574	0.670	1.953	0.316		
Quadratic	1.144	0.070	-1.039	0.200	1.818	0.450	0.777	0.084	-0.663	0.222	0.903	0.361
Genetic												
Constant	12.690	1.338										
Linear	-2.409	0.337	2.162	0.199								
Quadratic	-0.479	0.044	0.046	0.056	1.155	0.220						
Constant	11.048	1.056	-2.290	0.220	-0.467	0.045	11.225	2.913				
Linear	-1.778	0.173	1.011	0.237	-0.092	0.021	-1.864	0.479	0.727	0.238		
Quadratic	-0.261	0.026	0.301	0.070	0.296	0.101	-0.310	0.078	0.129	0.073	0.209	0.032

[1] HF = Holstein-Friesian.

[2] JE = Jersey.

Table 18. Variances and relative variances computed for a lactation of 270 DIM from the multi-breed model for purebred and crossbred animals

Breed composition		Phenotypic	Permanent environment		Additive genetic		Herd period		Residual	
HF[1]	JE[2]	Variance	Variance	Relative	Variance	Relative	Variance	Relative	Variance	Relative
1.00	0.00	35.6	17.0	0.48	10.2	0.29	8.42	0.24	0.04	0.00
0.75	0.25	29.4	13.9	0.47	9.58	0.33	5.84	0.20	0.04	0.00
0.50	0.50	24.2	11.2	0.46	9.15	0.38	3.82	0.16	0.04	0.00
0.25	0.75	20.1	8.79	0.44	8.91	0.44	2.36	0.12	0.04	0.00
0.00	1.00	17.0	6.71	0.39	8.85	0.52	1.45	0.09	0.04	0.00

[1] HF = Holstein-Friesian.

[2] JE = Jersey.

Only JE showed the large increase in h². Previous research by Lofgren *et al.* (1985) also showed a greater sensitivity of this breed, compared with HF, to changes in herd environments, expressed as means and standard deviations of contemporary groups adjusted for genetic differences. As shown in Table 13, means and standard deviations for daily milk yield were approximately 30% greater in the multi-breed data sets compared with the purebred JE data sets. Lofgren *et al.* (1985) reported that similar differences in means and standard deviations, but in a purebred setting, generated a substantial increase in h² (as much as 50%). Earlier studies (*e.g.* Legates, 1962, for fat yields) showed similar results, again in JE. It is very difficult to explain these results but they are in line with ours. It seems possible to imagine that, in particular, greater average yields in JE are linked to larger animals, which could express genetic differences. Because crossbred JE should be larger, this hypothesis could remain valid in our study.

From the (co)variances estimated with the multi-breed model, variances for milk as a function of DIM were computed for HF × JE crossbred animals. Results are given for first-cross animals of the 2 purebreds (HF= 50% and JE= 50%) and back-crosses between F1 and purebreds (HF = 75 or 25% and JE = 25 or 75%). Only phenotypic variances (Figure 5) as a function of DIM are presented, showing the evolution over the lactation. The variances for HF × JE crossbred animals were between those of the purebreds and followed the same trend as in purebred animals. These were confirmed by the values of variances for 270-day lactation yields presented in Table 18.

Correlations across breeds for the 3 Legendre polynomials are shown in Table 19. Virtually no differences existed among permanent environmental effects for both breeds. This could indicate that breed-specific permanent environmental effects are not important even if variance differences exist across breeds. Herd-period regression effects differed strongly, with correlations between 0.755 and -0.040. Because these effects translate into herd-specific lactation curves within breeds, these results could indicate not only that lactation curves among breeds are different, but also that breeds are managed differently, or at least that they react differently to a common management (Bryant *et al.*, 2007b). Direct interpretation of these results on a herd level could also indicate the existence of breed-specific competition effects that vary through the different lactation stages.

Table 19. Correlations among Legendre polynomial regressions across Holstein-Friesian and Jersey breeds estimated from the multi-breed model (values shown were obtained by back-transformation)

	Effect		
Polynomial	Herd period	Permanent environment	Additive genetic
Constant	0.755	0.989	0.926
Linear	0.616	0.974	0.807
Quadratic	-0.040	0.961	0.604

Correlations were very high for the constant genetic effect (0.926), but decreased to 0.604 for the quadratic genetic regressions. These results showed similar genetic rankings between the 2 breeds for the mean or overall milk yield. However, this does not mean that these ranking differences can be neglected completely, because top sires can rerank significantly, even with an overall correlation of 0.926. In addition, if other lactation shape parameters linked to persistency are considered, larger differences between HF and JE cattle seem to exist. This hypothesis is also supported by the values in Table 20 showing the genetic correlations between HF and JE cattle on a within lactation basis. Some correlations, especially those linking different DIM, were as low as 0.600. Phenotypic correlations are not given across breeds as they would not make much sense.

Because we used an animal model, most of the information on genetic correlations might come from the difference between dam-daughter regressions when daughters are purebred compared with crossbred. They also reflect expected ranking differences in purebred offspring compared with crossbred offspring of a purebred sire.

Heritability estimates as a function of DIM, as illustrated in Figure 7, were low at the beginning and the end of lactation, whereas they were greater towards mid-lactation for JE purebreds and crossbreds. The HF purebred animals had a lower h^2 than JE animals during lactation. Daily estimates of genetic and phenotypic correlations between different stages of lactation and daily estimates of heritabilities are presented in Table 20. Results were similar to those from the single-breed analysis, with correlations becoming weaker as the interval between tests increased. Genetic correlations were again greater for JE than for HF animals. However, this was not true for phenotypic correlations, for which values for JE were less.

Table 20. Heritabilities (diagonal), genetic correlations (above diagonal), and phenotypic correlations (below diagonal) for daily milk yield among first lactation estimated from multi-breed models for Holstein-Friesian and Jersey animals, genetic correlations among Holstein-Friesian and Jersey animals

	DIM					DIM				
DIM	5	60	120	180	270	5	60	120	180	270
	Holstein-Friesian					Holstein-Friesian × Jersey				
5	0.155	0.916	0.729	0.604	0.540	0.852	0.822	0.771	0.722	0.600
60	0.829	0.189	0.941	0.855	0.599	0.893	0.916	0.907	0.883	0.745
120	0.609	0.755	0.233	0.972	0.632	0.826	0.885	0.910	0.914	0.805
180	0.420	0.611	0.711	0.253	0.732	0.773	0.833	0.868	0.894	0.853
270	0.378	0.401	0.397	0.421	0.270	0.677	0.641	0.626	0.665	0.829
						Jersey				
5						0.218	0.981	0.943	0.918	0.845
60						0.719	0.297	0.989	0.971	0.855
120						0.537	0.625	0.331	0.993	0.865
180						0.391	0.515	0.575	0.314	0.907
270						0.347	0.384	0.397	0.413	0.281

CONCLUSIONS

Recently, an advanced test-day model that is adapted for its special breed structure was introduced in New Zealand. The present study investigated potential further advances for the future; (co)variance components and genetic parameters within and across breeds were estimated for HF and JE animals. Parameter estimates differed between single-breed and multi-breed analyses. However, this could also have been due to differences between the purebred and nonpurebred herds used for these analysis or could have been artefacts resulting from the choice of the models. This indicates limitations in the design of this study, because it was assumed a priori that purebred and nonpurebred herds were more similar, and the model was designed having this in mind. Future studies should consider our findings. The genetic correlations across breeds, computed from the multi-breed model, showed additive genetic differences, especially in linear and quadratic Legendre polynomials, which are linked to persistency. The results of this study showed that breed-dependent additive breeding values could be estimated by the proposed multi-breed model and could therefore provide a theoretically better tool to evaluate crossbred dairy cattle population as found in New Zealand. However, the situation presented here was simplified because only two breeds were analysed together. A routine model might require 10 or more different breed effects. Despite this, if in the future available computing resources increase and also if detecting genetic differences in lactation shape parameters becomes even more of an issue, models similar to the one presented in this study would be possible and worth considering. According to our findings, such a model could use a single permanent environmental effect. However, it would need distinct herd-period effects because result indicated the existence of breed-specific competition effects that varied through the different lactation stages. Scaling of variances would always be needed because we showed that rather large variance differences existed. Even intrabreed h^2 differences may exist according to the environment (purebred *vs* multi-breed herds). Therefore, future studies and a routine evaluation would still require the development of an improved methodology, and the model presented here is only a first step. Improved methodology would also address heterosis more correctly by estimating general, but potentially also specific, heterosis. Such a model would also be computationally very challenging because of the simultaneous presence of a larger number of breeds than the 2 breeds used in this study.

ACKNOWLEDGMENTS

Nicolas Gengler, who is Research Associate of the National Fund for Scientific Research (Brussels, Belgium), acknowledged its support. Additional support was provided through Grants F.4552.5 and 2.4507.02F (2) of the National Fund for Scientific Research. Manuscript review and useful comments by P. M. VanRaden (AIPL-ARS-USDA, Beltsville) and two anonymous reviewers are acknowledged.

LITERATURE CITED

Ahlborn G., and L. Dempfle. 1992. Genetic parameters for milk production and body size in New Zealand Holstein-Friesian and Jersey. Livest. Prod. Sci. 31:205-219

Bryant J. R., N. López-Villalobos, J. E. Pryce, C. W. Holmes, D. L. Johnson, and D. J. Garrick. 2007a. Environmental sensitivity in New Zealand dairy cattle. J Dairy Sci. 90:1538-1547.

Bryant J. R., N. López-Villalobos, J. E. Pryce, C. W. Holmes, D. L. Johnson, and D. J. Garrick. 2007b. Effect of environment on the expression of breed and heterosis effects for production traits. J Dairy Sci. 90:1548-1553.

Garrick, D.J., B.L. Harris, and D.L. Johnson. 1997. The across-breed evaluation of dairy cattle in New Zealand. Proc. Assoc. Advmt. Anim. Breed Genet. 12:611-615.

Harris B.L. 1994. Animal model for New Zealand dairy cattle evaluation. Proc. 1994 Interbull Meeting Ottawa, Canada. 10:1-13.

Harris, B.L. 1995. Animal evaluation methods for New Zealand dairy cattle. Livestock Improvement Corporation technical report. Hamilton, New Zealand. Pages 1-28.

Harris, B.L., and E.S. Kolver. 2001. Review of Holsteinization on intensive pastoral dairy farming in New Zealand. J. Dairy Sci. 84:E56-E61.

Harris, B. L., Winkelman, A. M., Johnson, D. L. and Montgomerie W. A. 2006. Development of a national production testday for New Zealand. Interbull Bull. 35:27-30

Interbull. 2007. Description of National Genetic Evaluation Systems for dairy cattle as applied in different Interbull member countries. http://www-interbull.slu.se/national_ges_info2/framesida-ges.htm. Accessed Oct. 19, 2007.

Jensen, J., E. A. Mäntysaari, P. Madsen, and R. Thompson. 1997. Residual maximum likelihood estimation of (co)variance components in multivariate mixed linear models using average information. J. Indian Soc. Agric. Stat. 49:215-236.

Johnson, D.L. 1996. Estimation of lactation yield from repeated measures of test day yields. Proc. New Zealand Soc. Anim. Prod. 56:16-18.

Legates, J.E. 1962. Heritability of fat yields in herds with different production levels. J. Dairy Sci. 45:990-993.

Lofgren, D.L., W.E. Vinson, R.E. Pearson, and R.L. Powell. 1985. Heritability of milk yield at different herd means and variance for production. J. Dairy Sci. 68:2737-2739.

Lopez-Villalobos, N., and D.J. Garrick. 1996. Profitability of rotational crossbreeding programmes in commercial dairy herds. Proc. New Zealand Soc. Anim. Prod. 56:216-220.

Lopez-Villalobos, N., and D.J. Garrick. 1997. Crossbreeding for increased profitability. Proc. Massey University Dairy Farmers Conference, New Zealand. 49:164-172.

Lopez-Villalobos, N., and D.J. Garrick. 2002. Economic heterosis and breed complementarity for dairy cattle in New Zealand. Proc. 7th World Congr. on Genet. Appl. Livest. Prod., Montpellier, France.

Lopez-Villalobos, N., D.J. Garrick, and C.W. Holmes. 2002. Genetic opportunities to improve milk value in New Zealand. Proc. New Zealand Soc. Anim. Prod. 62:90-94.

Lutaaya E., I. Misztal, J. W. Mabry, T. Short, H. H. Timm, and R. Holzbauer. 2001. Genetic parameter estimates from joint evaluation of purebreds and crossbreds in swine using the crossbred model J Anim Sci. 79:3002-3007.

Meyer, K. 1997. An ‘average information’ restricted maximum likelihood algorithm for estimating reduced rank genetic covariance matrices or covariance functions for animal models with equal design matrices. Genet. Sel. Evol. 29:97-116.

Misztal, I., T. Strabel, J. Jamrozik, E. A. Mäntysaari, and T.H.E. Meuwissen 2000. Strategies for estimating the parameters needed for different test-day models. J. Dairy Sci. 83:1125-1134.

Misztal, I. 2002. REMLF90 manual. Department of Animal and Dairy Science, University of Georgia, USA. Available: http://nce.ads.uga.edu/~ignacy/ Accessed Oct. 19, 2007.

Montgomerie, W.A. 2002. Experiences with dairy cattle crossbreeding in New Zealand. Pages 1-12. 53rd Annual Meeting of the European Association for Animal Production, Cairo, Egypt. Available: http://www.aeu.org.nz/news/EAAP_Xbreed_Sept2002.PDF Accessed Oct. 19, 2007.

Newman S., A. Reverter, and D. J. Johnston. 2002. Purebred-crossbred performance and genetic evaluation of postweaning growth and carcass traits in Bos indicus x Bos taurus crosses in Australia. J Anim Sci. 80:1801-1808.

Roso V. M., F. S. Schenkel, S. P. Miller, and J. W. Wilton. 2005. Additive, dominance, and epistatic loss effects on preweaning weight gain of crossbred beef cattle from different Bos taurus breeds. J Anim Sci. 83:1780-1787.

Touchberry, R.W. 1992. Crossbreeding effects in dairy cattle: the Illinois experiment, 1949 to 1969. J. Dairy Sci. 75:640-667.

VanRaden, P.M., and A.H. Sanders. 2001. Economic merit of crossbred and purebred US dairy cattle. J. Dairy Sci. 86:1036-1044.

Wei, M., and J.H.J van der Werf. 1995. Genetic correlation and heritabilities for purebred and crossbred performance in poultry egg production traits. J. Anim. Sci. 73:2220-2226.

Zumbach B., I. Misztal, S. Tsuruta, J. Holl, W. Herring, and T. Long. 2007. Genetic correlations between two strains of Durocs and crossbreds from differing production environments for slaughter traits. J Anim Sci. 85:901-908.

CHAPTER IV.2. USEFULNESS OF MULTI-BREED MODELS IN GENETIC EVALUATION OF DIRECT AND MATERNAL CALVING EASE IN HOLSTEIN AND BELGIAN BLUE WALLOON PUREBREDS AND CROSSBREDS

Based on: Vanderick, S., A. Gillon, G. Glorieux, M. Mayeres, R. R. Mota and N. Gengler. 2017. Estimation of genetic parameters for direct and maternal calving ease in Holstein and Belgian Blue Walloon cattle by using multi-breed models. Livest. Sci. 198:129-137.

CHAPTER IV.2. USEFULNESS OF MULTI-BREED MODELS IN GENETIC EVALUATION OF DIRECT AND MATERNAL CALVING EASE IN HOLSTEIN AND BELGIAN BLUE WALLOON PUREBREDS AND CROSSBREDS

OUTLINE

As previously stated calving ease is of great economic importance in dairy and beef cattle and should be considered in breeding programs. The development and implementation of the current genetic evaluation system of this maternally influenced trait for the Walloon Holstein dairy cattle have been presented in Chapters III.1 and III.2. A genetic evaluation system for calving ease also exists to predict genetic merit of Walloon dual-purpose Belgian Blue cattle. As these two systems are within breed systems, only data from purebred animals are used, thus ignoring data from crossbred animals. Chapter II indicated that there could be some benefits in performing a joint genetic evaluation of purebreds and crossbreds. A multi-breed model using breed-specific random regressions was developed in Chapter IV.1. Therefore, the objective of this chapter was to explore the feasibility and usefulness of a joint genetic evaluation for calving ease of Holstein and Belgian Blue Walloon cattle, using data from purebreds and crossbreds. Thus, based on results and statements of previous chapters as well as literature, single-breed and multi-breed linear animal models were adapted and assessed for the estimation of the genetic parameters for direct and maternal effects on calving ease. Then, the relative merits of these models were compared in terms of goodness of fits. The use and comparison of two multi-breed models using different functions of breed proportions as random regressions are what make the originality of this study.

ABSTRACT

The objective of this study was to verify the feasibility of a joint genetic evaluation system for calving ease trait of Belgian Blue (BBB) and Holstein (HOL) Walloon cattle based on data of purebred and crossbred animals. Variance components and derived genetic parameters for purebred BBB and HOL animals were estimated by using single-breed linear animal models. This analysis showed clear genetic differences between breeds. Estimates of direct and maternal heritabilities (± standard error) were 0.34 (±0.02) and 0.09 (±0.01) for BBB, respectively, but only 0.09 (±0.01) and 0.04 (±0.01) for HOL, respectively. Moreover, a significant negative genetic correlation between direct and maternal effects was obtained in both breeds: -0.46 (±0.04) for BBB and -0.29 (±0.11) for HOL. Variance components and derived genetic parameters for purebred BBB and HOL and crossbred BBB×HOL cattle were then estimated by using two multi-breed linear animal models: one based on a random regression test-day model (Model MBV), and one based on a random regression model (Model MBSM). Both multi-breed models use different functions of breed proportions as random regression predictors, thereby enabling modelling different additive effects according to animal's breed composition. The main difference between these models is the way in which relationships between breeds are accounted for in the genetic (co)variance structure. Genetic parameters differed between single-breed and multi-breed analysis, but are similar to the literature. For BBB, estimates of direct and maternal heritabilities (±SE) were 0.45 (±0.07) and 0.08 (±0.01) by using Model MBV, and 0.45 (±0.08) and 0.09 (±0.02) for Model MBSM, respectively. For HOL, these estimates were 0.18 (±0.04) and 0.05 (±0.01) using Model MBV, and 0.16 (±0.04) and 0.05 (±0.01) for Model MBSM, respectively. Reliability gains (up to 25%) indicated that the use of crossbred data in the multi-breed models had a positive influence on the estimation of genetic merit of purebred animals. A slight re-ranking of purebred sires and maternal grandsires was observed between single-breed and multi-breed models. Moreover, both multi-breed models can be considered as quasi-equivalent models because they performed almost equally well with respect to MSE and correlations, for purebred and crossbred animals.

Keywords: crossbreeding; covariance structure; genetic merit; reliability

INTRODUCTION

Crossbreeding is commonly used in beef cattle production systems and it is known to be an important tool to increase the efficiency of meat production through heterosis and breed complementarity. In dairy cattle, for several years, the fall in the prices of milk and dairy products coupled with growing health, fertility and calving performance concerns for almost all dairy breeds, have driven dairy breeders and producers to increasingly adopt crossbreeding in their herds (Sørensen *et al.*, 2008; Fouz *et al.*, 2013). Additionally, increased inbreeding rate and homozygosity in those breeds have also contributed to the growing interest in crossbreeding around the world (Hansen, 2006; Sørensen *et al.*, 2008). Two crossbreeding strategies are mainly used to exploit heterosis in a commercial dairy herd. First, the rotational crossbreeding program which includes two or more dairy breeds being used alternately: most dairy cattle crossbreeding systems use three breeds to optimize the average level of heterosis across generations (Hansen, 2006). The second strategy is to mate dairy or dual-purpose cows with beef bulls, adopting a simple two-breed cross on animals not used to produce purebred replacements in dairy farms. The final cross, known as “industrial or terminal cross”, is sold to commercial producers with higher price than selling purebred animals (Sørensen *et al.*, 2008; Fouz *et al.*, 2013).

The ability to calve easily is an economically important trait in dairy and beef cattle production systems. This trait affects the profitability of herds, animal welfare and thereby acceptability of the production systems by the consumer (Mee, 2008; Barrier *et al.*, 2013). Thus, selection to improve calving ease is an important breeding goal that may grow in importance over time.

Currently, two distinct genetic evaluations are carried out for calving ease in the Walloon Region of Belgium: one for the Holstein (HOL) dairy population (Vanderick *et al.*, 2013); and another for the dual-purpose Belgian Blue population. However, data from crossbred descendants of artificial insemination (AI) bulls of Belgian Blue breed (BBB) and HOL cows are currently not used in either one of them. This is because these genetic evaluations compare animals only within breed, based on purebred information, therefore excluding information from crossbred animals. The BBB sires of these crossbreds are from two sources, primarily from the beef (meat) production population but also from the dual-purpose population (as determined from the Walloon cattle pedigree consulted in April 2016).

Originally, the BBB was a dual-purpose breed combining dairy and beef production. From the 1950s to the 1980s, the selection process was mainly focused on meat production and selection for milk production discontinued leading to the current beef type (*i.e.* beef animals with a muscular hypertrophy). However, some breeders did not adopt this breeding goal and therefore, currently a small population of the original dual-purpose BBB remained. Thus, the BBB breed is composed of two strains; the most important beef-purpose strain in a suckler herd system, and the dual-purpose one, with currently about 4000 milking cows recorded in the Walloon milk database consulted in April 2016.

A study by Vanderick *et al.* (2014) estimated genetic parameters for direct and maternal calving ease for Walloon HOL dairy cattle. Direct and maternal heritability (h^2) estimates for calving ease ranged from 0.07 to 0.12 and from 0.02 to 0.04, respectively, depending on the model fitted (linear *vs* threshold). These results were in the range of previously published h^2 estimates of calving ease in HOL dairy cattle (*e.g.* López de Maturana *et al.*, 2007; Eaglen *et al.*, 2012). The current genetic evaluation system for calving ease in Walloon HOL dairy cattle is based on results of Vanderick *et al.* (2013).

The current genetic evaluation system for calving ease for Walloon dual-purpose Belgian Blue is based on the methodology proposed by Mayeres *et al.* (2007). This non-published study estimated direct and maternal heritabilities around 0.34 and 0.11, respectively, and their correlation to be approximately -0.66. These results were consistent with most values reported in the literature for other dual-purpose and beef breeds (*e.g.* Hagger and Hofer, 1990; Bennett and Gregory, 2001).

Therefore, given the obvious differences between h^2 estimates of HOL and Dual-purpose Belgian Blue populations, a joint genetic evaluation for calving ease of these populations requires strategies that allow the use of a heterogeneous genetic covariance structure in order to keep the genetic features of each population. Several multi-breed models using heterogeneous (co)variances across breeds have been proposed (Lo *et al.*, 1993; García-Cortés and Toro, 2006; Vanderick *et al.*, 2009; Strandén and Mäntysaari, 2013). These multi-breed models allow purebred and crossbred data to be combined and to further perform joint genetic evaluations of purebred and crossbred animals. Toward this end, the main goal of this study was to verify the feasibility of a joint genetic evaluation for calving ease in Walloon BBB and HOL animals. To achieve this, 1) genetic parameters for direct and maternal effects on calving ease of BBB and HOL cattle in the Walloon Region of Belgium were estimated using single-breed linear animal models; 2) two literature-based multi-breed linear animal models (Vanderick *et al.*, 2009; Strandén and Mäntysaari, 2013) were adapted and then assessed for the estimation of the breed-specific additive genetic parameters for direct and maternal for calving ease; 3) the relative merits of these single-breed and multi-breed models were compared in terms of goodness of fit.

MATERIALS AND METHODS

PHENOTYPIC DATA

In the Walloon Region of Belgium, calving ease in Walloon BBB and HOL cows is recorded by breeders on a voluntary basis at birth registration, and collected by the Walloon Breeding Association (awé). In this study, calving ease was subjectively scored by breeders on a scale from 1 to 4 as follows: 1) Caesarean section and embryotomy, 2) hard pull, 3) easy pull (slight assistance), and 4) normal (no help) (Vanderick *et al.*, 2014). Records from calves born between 2000 and 2015 were used. Data editing was largely based on that previously described by Vanderick *et al.* (2014). Breed proportions were computed for each calf and its parents from a pedigree file of 6 346 577 animals of cattle breeds raised in the Walloon Region of Belgium. Only animals of BBB and HOL inheritance were kept to estimate genetic parameters. In the context of this study, purebreds were animals with at least 95% of BBB or HOL genes, therefore, only calves in which the sum of the two breed proportions was equal to 95% or more remained. It should be emphasised that no distinction was made between the two strains of BBB, *i.e.* between the beef-purpose and the dual-purpose. Thus, the edited data set consisted of 427 202 calving records from 1670 Walloon herds: 55% of herds were composed of BBB purebreds, 23% were composed of HOL purebreds and 22% by BBB × HOL crossbreds as well as by BBB and HOL purebreds. Similar to purebred animals, purebred herds in this study were considered as herds with at least 95% of average breed compositions of BBB or HOL genes. Crossbreds were animals, in this study, with at least 50% but less than 95% BBB or HOL genes.

From this edited data set, two data sets were prepared. Within each purebred herd, animals with less than 95% of BBB or HOL genes were removed. In order to remove BBB purebred herds using caesarean section (CS) as routine management tool (Kolkman *et al.*, 2007), BBB herds were selected according to their caesarean mean rate. Thus, BBB purebred herds with at least 3% of calvings without CS were kept. The first data set (data set I) included records from 55 319 BBB purebred calves born in 257 Walloon BBB purebred herds from 2401 sires, 30 074 dams and 2258 maternal grandsires. Pedigree was traced back to animals born in 1985 and hence the pedigree file for data set I consisted of 102 221 records. The second data set (data set II) comprised records from 40 535 HOL purebred calves originating from 383 Walloon HOL purebred herds, from 2700 sires, 27 194 dams and 2379 maternal grandsires. The total number of animals in the pedigree was 102 820.

In addition, a third data set (data set III) was created by combining calving scores from crossbred calves extracted from the edited data set with purebred calving scores from data sets I and II. Data set III included records from 100 757 purebred and crossbred calves distributed as follows: 55% of BBB purebred, 40% of HOL purebred, and 5% of crossbred calves. More than half of crossbreds were 50% BBB × 50% HOL. The total number of animals including ancestors without records was 219 625.

Descriptive statistics of these three data sets are presented in Table 21. Calving ages of dam were grouped as 18-24, 25-26, 27-28, 29-30, 31-33, 34-37, 38-41, 42-45, 46-49, 50-56, 57-65, 66-79 and more than 79 months at calving (*i.e.* 13 classes) for all data sets. Four seasons of calving were used in the analysis: winter (January – March), spring (April – June), summer (July – September) and autumn (October – December). Scores were not transformed because the use of equally spaced scores has been suggested to have negligible effects on h^2 estimates (Abdel-Azim and Berger, 1999).

Table 21: Descriptive statistics of the data sets I, II and III

Item	Data set I (BBB[a] purebreds)		Data set II (HOL[a] purebreds)		Data set III (Data set I + Data set II + crossbred animals)	
	No. of observations	Percentage	No. of observations	Percentage	No. of observations	Percentage
Number of animals	55 319		40 535		100 757	
Female calves	31 808	57.5	28 415	70.1	62 562	62.1
Male calves	23 511	42.5	12 120	29.9	38 195	37.9
Final pedigree file	102 221		102 820		219 625	
Herds	257		383		922	
Herd × year of calving	2438		3337		6718	
With only calving by CS[b]	1180	48.4	8	0.2	1424	21.2
Without calving by CS[b]	115	4.7	3021	90.5	3618	53.9
Sires with progeny records	2401		2700		5595	
Dams	30 074		27 194		60 774	
Maternal grandsires	2258		2379		4984	
Maternal granddams	20 999		22 055		46 109	
Calving ease						
1. Caesarean and embryotomy	46 248	83.6	446	1.1	48 346	47.9
2. Hard pull	940	1.7	2189	5.4	3384	3.4
3. Easy pull	2212	4.0	10 296	25.4	13 298	13.2
4. Normal	5919	10.7	27 604	68.1	35 729	35.5

[a] BBB = Belgian Blue breed, HOL= Holstein breed

[b] CS= Caesarean section

MODELS OF ANALYSIS

A previous study provided evidence that there was no clear advantage to use a threshold animal model rather than a linear animal model to evaluate calving ease for Walloon HOL dairy cattle (Vanderick *et al.*, 2014). Therefore, only linear animal models were considered in the present study to analyse calving ease of BBB and HOL cattle in the Walloon Region of Belgium.

Firstly, a single-breed linear animal model was used to estimate (co)variance components and derive genetic parameters for purebred BBB and HOL cattle by using data sets I and II, respectively. Hereafter, this study was called a single-breed analysis. Secondly, the data set III was used to estimate (co)variance components and derive genetic parameters for purebred BBB and HOL as well as crossbred BBB×HOL cattle by using two multi-breed linear animal models. The methodology of the first multi-breed model was derived from the random regression test-day model proposed by Vanderick *et al.* in 2009 and the methodology of the second multi-breed model was based on the random regression multi-breed model proposed by Strandén and Mäntysaari in 2013, considered as an approximation of the split multi-breed model described by García-Cortés and Toro (2006). Both multi-breed models use breed proportions as random regression predictors, thereby enabling modelling different additive effects according to breed composition of each animal. Thus, these multi-breed models generate as many variances and estimated breeding values (EBV) as the number of breeds in the data set. Hereafter, this study was called a multi-breed analysis. This strategy of single-breed and multi-breed analyses was used to simplify actual estimations by estimating within and then across breeds. Moreover, this strategy allowed us to compare the results from the 2 analyses.

Single-Breed Analysis.

(Co)variance components were estimated for the BBB and HOL breeds separately by using data sets I and II, respectively. Calving ease was fitted as a continuous trait by using the univariate linear animal model described in Vanderick *et al.* (2014). The single-breed model used in this analysis included fixed effects of season, herd and combined effects of sex of calf × age of dam classes × group of parities (two groups: 1) 1st parity and 2) from the 2nd to the 5th parity). Direct and maternal additive genetic effects, maternal permanent environmental effects, herd × year of calving effects, which were included to account for the variability in the frequency of dystocia among herds and years within herds, and residual were fitted as random effects. Model designated as Model L1 was fitted with an estimated genetic correlation between direct and maternal additive genetic effects and model designated as Model L2 with a genetic correlation between genetic effects equal to zero.

Multi-Breed Analysis

The following fixed effects were included in both investigated multi-breed models: season effects, combined effects of herd by breed-of-dam (two groups: HOL purebred or crossbred dams and BBB purebred dams) and combined effects of sex of calf × age of dam classes × group of parities (two groups as described in single-breed analysis) × breed-of-calf (two management groups: 1) HOL purebred or crossbred calf and 2) BBB purebred calf). These groups of breed-of-dam and breed-of-calf were included to account for breed-of-dam and breed-of-calf differences, but also to account for differences between purebred and crossbred calvings. Calvings from HOL purebred or crossbred dams (calves) were considered as being similar from a management perspective. Furthermore, covariates for direct and maternal heterosis coefficients as well as for direct and maternal recombination loss coefficients were included in both multi-breed models. The heterosis and recombination loss coefficients were calculated for each calf and for each dam from data set III as $1-\sum_{X=1}^{n} f_{sire}^{(X)} \times f_{dam}^{(X)}$ and $1-\sum_{X=1}^{n} \frac{\left(f_{sire}^{(X)}\right)^2 + \left(f_{dam}^{(X)}\right)^2}{2}$ respectively, in which $f_{sire}^{(X)}$ and $f_{dam}^{(X)}$ are the proportion of breed X with $X = \{\text{BBB,HOL}\}$; in the sire and dam, respectively (VanRaden and Sanders, 2003).

The first multi-breed model (hereafter called Model MBV) was based on the random regression test-day animal model developed by Vanderick *et al.* (2009). This multi-breed model separates the random effects by breed and these parts are correlated in the variance structure. Therefore, correlation between breeds is taken into account in the estimation of each random effect. Thus, in this study, the direct or maternal additive value of an animal is decomposed into several genetic parts depending on their breed of origin. For the purpose of this study, this multi-breed model was adapted and hence, Model MBV can be written as:

$$\begin{aligned}\mathbf{y}_{CE} &= \mathbf{X\beta} + \mathbf{Z}_h^{(BBB)}\mathbf{h}^{(BBB)} + \mathbf{Z}_p^{(BBB)}\mathbf{p}^{(BBB)} + \mathbf{Z}_a^{(BBB)}\mathbf{a}^{(BBB)} + \mathbf{Z}_m^{(BBB)}\mathbf{m}^{(BBB)} \\ &+ \mathbf{Z}_h^{(HOL)}\mathbf{h}^{(HOL)} + \mathbf{Z}_p^{(HOL)}\mathbf{p}^{(HOL)} + \mathbf{Z}_a^{(HOL)}\mathbf{a}^{(HOL)} + \mathbf{Z}_m^{(HOL)}\mathbf{m}^{(HOL)} + \mathbf{e}\end{aligned} \quad (53)$$

where $\mathbf{y}_{CE}$ is a vector of observed calving ease scores, $\boldsymbol{\beta}$ is a vector of fixed effects and covariates; $\mathbf{h}^{(BBB)}$ ($\mathbf{h}^{(HOL)}$) is a vector of herd × year of calving random regression coefficients for BBB (HOL) breed; $\mathbf{p}^{(BBB)}$ ($\mathbf{p}^{(HOL)}$) is a vector of permanent maternal environmental random regression coefficients for BBB (HOL) breed; $\mathbf{a}^{(BBB)}$ ($\mathbf{a}^{(HOL)}$) is a vector of direct additive genetic random regression coefficients for BBB (HOL) breed; $\mathbf{m}^{(BBB)}$ ($\mathbf{m}^{(HOL)}$) is a vector of maternal additive genetic random regression coefficients for BBB (HOL) breed; $\mathbf{X}$ is the incidence matrix linking observations with fixed effects; $\mathbf{e}$ is a vector of residual effects; $\mathbf{Z}_h^{(BBB)}$ ($\mathbf{Z}_h^{(HOL)}$), $\mathbf{Z}_p^{(BBB)}$ ($\mathbf{Z}_p^{(HOL)}$), $\mathbf{Z}_a^{(BBB)}$ ($\mathbf{Z}_a^{(HOL)}$) and $\mathbf{Z}_m^{(BBB)}$ ($\mathbf{Z}_m^{(HOL)}$) are the breed-specific regressors. These breed-specific regressors are defined as $\mathbf{Z}_h^{(BBB)} = \mathbf{\Phi}_{calf}^{(BBB)}\mathbf{Z}_h$, $\mathbf{Z}_p^{(BBB)} = \mathbf{\Phi}_{dam}^{(BBB)}\mathbf{Z}_p$, $\mathbf{Z}_{\mathbf{a}}^{(BBB)} = \mathbf{\Phi}_{calf}^{(BBB)}\mathbf{Z}_a$, and $\mathbf{Z}_m^{(BBB)} = \mathbf{\Phi}_{dam}^{(BBB)}\mathbf{Z}_m$ for BBB and $\mathbf{Z}_h^{(HOL)} = \mathbf{\Phi}_{calf}^{(HOL)}\mathbf{Z}_h$, $\mathbf{Z}_p^{(HOL)} = \mathbf{\Phi}_{dam}^{(HOL)}\mathbf{Z}_p$, $\mathbf{Z}_a^{(HOL)} = \mathbf{\Phi}_{calf}^{(HOL)}\mathbf{Z}_a$, and $\mathbf{Z}_m^{(HOL)} = \mathbf{\Phi}_{dam}^{(HOL)}\mathbf{Z}_m$ for HOL where $\mathbf{\Phi}_{calf}^{(BBB)}$, $\mathbf{\Phi}_{dam}^{(BBB)}$, $\mathbf{\Phi}_{calf}^{(HOL)}$ and $\mathbf{\Phi}_{dam}^{(HOL)}$ are diagonal matrices containing the breed proportions of calf $\left(f_{calf}^{(BBB)}\text{ and } f_{calf}^{(HOL)}\right)$ and dam $\left(f_{dam}^{(BBB)}\text{ and } f_{dam}^{(HOL)}\right)$ for BBB and HOL, respectively; and $\mathbf{Z}_h$, $\mathbf{Z}_a$, $\mathbf{Z}_m$ and $\mathbf{Z}_p$ are the incidence matrices linking observations with respective effects. The residual effects were assumed independently distributed with the same residual variance σ_e^2.

The associated (co)variance matrices among the breed-specific random effects could then be shown grouped by type of effects:

$$Var\begin{bmatrix}\mathbf{h}^{(BBB)}\\ \mathbf{h}^{(HOL)}\end{bmatrix}=\begin{bmatrix}\left(\sigma_h^2\right)^{(BBB)} & \left(\sigma_h\right)^{(BBB\times HOL)}\\ \left(\sigma_h\right)^{(BBB\times HOL)} & \left(\sigma_h^2\right)^{(HOL)}\end{bmatrix}\otimes\mathbf{I}_h,$$

$$Var\begin{bmatrix}\mathbf{p}^{(BBB)}\\ \mathbf{p}^{(HOL)}\end{bmatrix}=\begin{bmatrix}\left(\sigma_p^2\right)^{(BBB)} & \left(\sigma_p\right)^{(BBB\times HOL)}\\ \left(\sigma_p\right)^{(BBB\times HOL)} & \left(\sigma_p^2\right)^{(HOL)}\end{bmatrix}\otimes\mathbf{I}_p, \text{ and}$$

$$Var\begin{bmatrix}\mathbf{a}^{(BBB)}\\ \mathbf{m}^{(BBB)}\\ \mathbf{a}^{(HOL)}\\ \mathbf{m}^{(HOL)}\end{bmatrix}=\begin{bmatrix}\left(\sigma_a^2\right)^{(BBB)} & \left(\sigma_{am}\right)^{(BBB)} & \left(\sigma_a^2\right)^{(BBB\times HOL)} & \left(\sigma_{am}\right)^{(BBB\times HOL)}\\ \left(\sigma_{am}\right)^{(BBB)} & \left(\sigma_m^2\right)^{(BBB)} & \left(\sigma_{ma}\right)^{(BBB\times HOL)} & \left(\sigma_m^2\right)^{(BBB\times HOL)}\\ \left(\sigma_a^2\right)^{(BBB\times HOL)} & \left(\sigma_{ma}\right)^{(BBB\times HOL)} & \left(\sigma_a^2\right)^{(HOL)} & \left(\sigma_{am}\right)^{(HOL)}\\ \left(\sigma_{am}\right)^{(BBB\times HOL)} & \left(\sigma_m^2\right)^{(BBB\times HOL)} & \left(\sigma_{am}\right)^{(HOL)} & \left(\sigma_m^2\right)^{(HOL)}\end{bmatrix}\otimes\mathbf{A}$$

where $\left(\sigma_h^2\right)^{(X)}$ is the variance of herd × year of calving effects, $\left(\sigma_p^2\right)^{(X)}$ is the variance of permanent maternal environmental effects, $\left(\sigma_a^2\right)^{(X)}$ is the variance of direct additive genetic effects, $\left(\sigma_m^2\right)^{(X)}$ is the variance of maternal additive genetic effects, and $\left(\sigma_{am}\right)^{(X)}$ is the direct-maternal additive covariance for breed X with $X=\{BBB,HOL\}$; $\left(\sigma_h\right)^{(BBB\times HOL)}$ is the herd × year of calving covariance between the BBB and HOL breeds; $\left(\sigma_p\right)^{(BBB\times HOL)}$ is the permanent maternal environmental covariance between the BBB and HOL breeds; $\left(\sigma_a^2\right)^{(BBB\times HOL)}$ and $\left(\sigma_m^2\right)^{(BBB\times HOL)}$ are the direct and maternal additive genetic variances between the BBB and HOL breeds, respectively; $\left(\sigma_{am}\right)^{(BBB\times HOL)}$ is the covariance between direct additive genetic effects for the BBB breed and maternal additive genetic effects for the HOL breed; $\left(\sigma_{ma}\right)^{(BBB\times HOL)}$ is the covariance between maternal additive genetic effects for the BBB breed and direct additive genetic effects for the HOL breed; $\mathbf{I}_h$ and $\mathbf{I}_p$ are identity matrices; and $\mathbf{A}$ is the additive relationship matrix. Covariances between genetic and environmental effects were assumed to be zero, and no variances due to dominance or epistatic effects were assumed to exist.

The second multi-breed model (hereafter called Model MBSM) was based on the random regression model presented by Strandén and Mäntysaari (2013), considered as an approximation of the split multi-breed model proposed by García-Cortés and Toro (2006). In the split multi-breed model, the additive value is split into several independent parts depending upon their genetic origin; the variance-covariance structure of each part is determined by a partial relationship matrix, which describes relationships only according to genetic origin. The algorithms to construct directly the inverse of these partial matrices, by using breed composition, were provided by García-Cortés and Toro (2006). To relieve the need of build these partial relationship matrices, an equivalent random regression multi-breed model can be used (Strandén and Mäntysaari, 2013). For the purpose of this study, the model proposed by Strandén and Mäntysaari (2013) was adapted and hence, Model MBSM can be written as follows:

$$\begin{aligned}\mathbf{y}_{CE} &= \mathbf{X\beta} + \mathbf{Z}_h^{(BBB)}\mathbf{h}^{(BBB)} + \mathbf{Z}_p^{(BBB)}\mathbf{p}^{(BBB)} + \mathbf{W}_a^{(BBB)}\mathbf{a}^{(BBB)} + \mathbf{W}_m^{(BBB)}\mathbf{m}^{(BBB)} \\ &+ \mathbf{Z}_h^{(HOL)}\mathbf{h}^{(HOL)} + \mathbf{Z}_p^{(HOL)}\mathbf{p}^{(HOL)} + \mathbf{W}_a^{(HOL)}\mathbf{a}^{(HOL)} + \mathbf{W}_m^{(HOL)}\mathbf{m}^{(HOL)} \qquad (54) \\ &+ \mathbf{W}_a^{(Seg)}\mathbf{a}^{(Seg)} + \mathbf{W}_m^{(Seg)}\mathbf{m}^{(Seg)} + \mathbf{e}\end{aligned}$$

Several features distinguished Model MBSM from Model MBV. The first difference was in the definition of the breed-specific regressors for the additive genetic random effects: $\mathbf{W}_a^{(BBB)} = \mathbf{\Psi}_{calf}^{(BBB)}\mathbf{Z}_a$ and $\mathbf{W}_m^{(BBB)} = \mathbf{\Psi}_{dam}^{(BBB)}\mathbf{Z}_m$ for BBB, and $\mathbf{W}_a^{(HOL)} = \mathbf{\Psi}_{calf}^{(HOL)}\mathbf{Z}_a$, and $\mathbf{W}_m^{(HOL)} = \mathbf{\Psi}_{dam}^{(HOL)}\mathbf{Z}_m$ for HOL where $\mathbf{\Psi}_{calf}^{(BBB)}$, $\mathbf{\Psi}_{dam}^{(BBB)}$, $\mathbf{\Psi}_{calf}^{(HOL)}$ and $\mathbf{\Psi}_{dam}^{(HOL)}$ are diagonal matrices containing the square root of breed proportions of calf $\left(\sqrt{f_{calf}^{(BBB)}} \text{ and } \sqrt{f_{calf}^{(HOL)}}\right)$ and dam $\left(\sqrt{f_{dam}^{(BBB)}} \text{ and } \sqrt{f_{dam}^{(HOL)}}\right)$ for BBB and HOL, respectively. The square root of breed proportions was used to equalize the proportion of direct and maternal additive genetic variances accounted for by breeds and avoid high variation between purebred and crossbred direct and maternal additive genetic variances when fitting $f_{calf}^{(X)}$ and $f_{dam}^{(X)}$, respectively, for breed X with $X = \{BBB, HOL\}$. Secondly, an extra additive genetic component due to the difference in allelic frequencies between the two breeds was fitted in Model MBSM. This component is called the segregation term and is equal to the difference in additive variances between second-cross (F2) and first-cross generation (F1) (Lo *et al.*, 1993). Thus, $\mathbf{W}_a^{(Seg)}\mathbf{a}^{(Seg)}$ and $\mathbf{W}_m^{(Seg)}\mathbf{m}^{(Seg)}$ where $\mathbf{a}^{(Seg)}$ and $\mathbf{m}^{(Seg)}$ are the vectors of direct and maternal genetic random regression effects for breed segregation, respectively; and $\mathbf{W}_a^{(Seg)}$ and $\mathbf{W}_m^{(Seg)}$ are the breed segregation specific regressions defined as $\mathbf{W}_a^{(Seg)} = \mathbf{H}_{calf}\mathbf{Z}_a$ and $\mathbf{W}_m^{(Seg)} = \mathbf{H}_{dam}\mathbf{Z}_m$ where $\mathbf{H}_{calf}$ and $\mathbf{H}_{dam}$ are diagonal matrices containing the square root of breed segregation proportion, for calf and dam, respectively. The breed segregation proportion for an animal *i* was computed as follows:

$$h_i^{(BBB-HOL)} = 2\left(f_{sire}^{(BBB)} f_{sire}^{(HOL)} + f_{dam}^{(BBB)} f_{dam}^{(HOL)}\right) \quad (55)$$

where $f_{sire}^{(BBB)}$ and $f_{sire}^{(HOL)}$ are the proportions of genes of its sire coming from the BBB and HOL breed, respectively; and $f_{dam}^{(BBB)}$ and $f_{dam}^{(HOL)}$ are the proportions of genes of its dam coming from the BBB and HOL breed, respectively. Finally, covariances among breed-specific additive genetic effects were not accounted for in Model MBSM and were replaced by the segregation variances for the direct and maternal effects ($\sigma_a^{2(Seg)}$ and $\sigma_m^{2(Seg)}$), respectively. Thus, contrary to Model MBV, the different parts of additive value are uncorrelated in the variance structure and the additive (co)variance matrices can be written as:

$$Var\begin{bmatrix} \mathbf{a}^{(BBB)} \\ \mathbf{m}^{(BBB)} \\ \mathbf{a}^{(HOL)} \\ \mathbf{m}^{(HOL)} \\ \mathbf{a}^{(Seg)} \\ \mathbf{m}^{(Seg)} \end{bmatrix} = \begin{bmatrix} \left(\sigma_a^2\right)^{(BBB)} & \left(\sigma_{am}\right)^{(BBB)} & 0 & 0 & 0 & 0 \\ \left(\sigma_{am}\right)^{(BBB)} & \left(\sigma_m^2\right)^{(BBB)} & 0 & 0 & 0 & 0 \\ 0 & 0 & \left(\sigma_a^2\right)^{(HOL)} & \left(\sigma_{am}\right)^{(HOL)} & 0 & 0 \\ 0 & 0 & \left(\sigma_{am}\right)^{(HOL)} & \left(\sigma_m^2\right)^{(HOL)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \left(\sigma_a^2\right)^{(Seg)} & \left(\sigma_{am}\right)^{(Seg)} \\ 0 & 0 & 0 & 0 & \left(\sigma_{am}\right)^{(Seg)} & \left(\sigma_m^2\right)^{(Seg)} \end{bmatrix} \otimes \mathbf{A}$$

where $\left(\sigma_a^2\right)^{(Seg)}$ and $\left(\sigma_m^2\right)^{(Seg)}$ are the direct and maternal segregation variances between breeds BBB and HOL; and $\left(\sigma_{am}\right)^{(Seg)}$ is the direct-maternal genetic segregation covariance. The associated (co)variance matrices among the breed-specific herd × year of calving random regression effects and among the breed-specific permanent maternal environmental random regression effects are identical to those from Model MBV. Furthermore, as for Model MBV, covariances between genetic and environmental effects were assumed to be zero, and no variances due to dominance or epistatic effects were assumed to exist.

ESTIMATION OF (CO)VARIANCE COMPONENTS AND GENETIC PARAMETERS

For all tested models, (co)variance components were estimated using expectation maximization REML (Misztal *et al.*, 2014). Standard errors (SE) of (co)variance estimates were obtained using average information REML (Misztal *et al.*, 2014).

Genetic parameters direct and maternal h² were defined across herds and computed as $h_d^2 = \sigma_a^2 \big/ \left(\sigma_a^2 + \sigma_{am} + \sigma_m^2 + \sigma_h^2 + \sigma_p^2 + \sigma_e^2\right)$ and $h_m^2 = \sigma_m^2 \big/ \left(\sigma_a^2 + \sigma_{am} + \sigma_m^2 + \sigma_h^2 + \sigma_p^2 + \sigma_e^2\right)$, respectively. Genetic correlations between direct and maternal effects were estimated by: $r_u\left(a,m\right) = \sigma_{am} \big/ \left(\sqrt{\sigma_a^2 * \sigma_m^2}\right)$. Approximated SE for genetic correlations and heritabilities were calculated using equations provided by Klei and Tsuruta (2008).

MODEL COMPARISON

Reliabilities and Rank Correlations

The comparison between single-breed and multi-breed models was based on reliabilities and Spearman’s rank correlations of direct and maternal EBV for sires and maternal grandsires with progeny. Reliability of direct and maternal EBV was defined as the squared correlation between true and predicted EBV and can be calculated as

$$r_{ixz}^2 = 1 - \left(pev_{ixz} / \sigma_{xz}^2\right) \qquad (56)$$

where r_{ixz}^2 is the reliability for animal *i* and breed *x* for genetic component *z* (*i.e.* direct or maternal effect), pev_{ixz} is the corresponding prediction error variance given by direct inversion of the coefficient matrix, and σ_{xz}^2 is the additive variance for breed *x* and for genetic component *z*.

Mean Squared Error and Pearson's Correlation

Moreover, the comparison between these models were also based on their goodness of fit and, to that end, two different criteria were computed: the mean squared error (MSE) and the Pearson's correlation between observed and predicted observations. The MSE was computed as

$$\mathrm{MSE} = \frac{1}{\mathrm{n}} \sum_{i=1}^{n} (y_{\mathrm{CE}i} - \hat{y}_{\mathrm{CE}i})^2 \qquad (57)$$

where y_{CE} and $\hat{y}_{\mathrm{CE}}$ correspond to the observed and predicted calving ease scores, respectively; n is the number of data points in data set.

The EBV and predictions of calving ease were computed with a BLUP approach for each of the following models: Model L1 used with the BBB and HOL purebred data sets (*i.e.* data set I and data set II, respectively) and both multi-breed models (*i.e.* Model MBV and Model MBSM) used with the full data set III.

RESULTS AND DISCUSSION

The descriptive statistics of the data sets I, II and III are presented in Table 21. Unbalanced sex ratios were observed in records, there were fewer male calves than females in the three data sets. The most unbalanced data set was data set II that contained calving records from HOL purebred herds. As suggested in a previous study for the calving ease trait in Walloon dairy cattle (Vanderick *et al.*, 2014), this imbalance may be due to the difference in economic value between male and female dairy calves, so that breeders prefer to record female over male calves. This recording bias in the number of female and male calves can lead to an underreporting of difficulty to calve (score < 4) as the calving of males is known to be more difficult in both beef and dairy cattle (*e.g.* Heins *et al.*, 2006; Fouz *et al.*, 2013).

For the single-breed analysis, calving records from 55 319 BBB purebred calves distributed over 2438 herd × year of calving combinations (data set I) and calving records from 40 535 HOL purebred calves distributed over 3337 herd × year of calving combinations (data set II) were analysed. As expected, most BBB purebred calves were delivered by CS (1). Nonetheless, about 15% of BBB purebred calves did not need any help (4) or only need of slight assistance (3). Most calvings of HOL purebred calves were classified as normal (4) or with slight assistance (3). Although around 84% of BBB purebred calves were born by CS, less than half of all combinations of herd × year of calving were exclusively composed of calvings by CS in data set I. Furthermore in this data set, about 5% of herd × year of calving classes were composed of calvings without CS. Distribution of herd × year of calving combinations in data set II is also provided in Table 21.

For the multi-breed analysis, calving records from 4903 crossbred calves were added to the almost 96 000 calvings from purebred data sets. This low number of crossbred calving records can be explained by a lack of economic incentives to encourage breeders and producers to register their crossbred animals.

SINGLE-BREED ANALYSIS

For single-breed analysis, (co)variances and derived genetic parameters of calving ease were estimated using linear animal models (Model L1 and Model L2) where each breed was analysed separately. Table 22 presents the estimates of (co)variance components and genetic parameters with their SE, and AIC values for Model L1 and Model L2. Standard errors were relatively small for all estimates. Estimates of (co)variance components for HOL purebred animals and their related genetic parameters agreed with those reported in a previous study on the same dairy cattle population and linear animal models (Vanderick *et al.*, 2014); the exception was the significant negative genetic correlation observed in the present study between direct and maternal effects for Model L1.

Concerning the genetic parameters of BBB purebred animals, they were consistent with those reported in the literature for calving ease in beef cattle and dual-purpose cattle (Hagger and Hofer, 1990; Koots *et al.*, 1994a; Phocas and Laloë, 2003; Berry and Evans, 2014). Note that the direct additive h² in Table 22 was higher than h² values often reported in literature for other beef breeds. However, Bennett and Gregory (2001) reported quite higher heritabilities for calving ease in beef and dual-purpose heifers (*i.e.* estimates around 0.40 and 0.20 for direct and maternal h², respectively). Moreover, in a non-published study on the birth traits in Walloon dual-purpose Belgian Blue population, a direct h² of 0.34 was estimated for calving ease (Mayeres *et al.*, 2007).

For both breeds, the additive genetic variance due to direct effects was greater than that due to maternal effects. Direct heritabilities were approximately four to five times as large as maternal heritabilities for BBB purebred animals and from two to three times for HOL purebred animals (Table 22). Greater estimates of direct compared with maternal heritabilities have also been found in previous studies in beef and dairy cattle (López de Maturana *et al.*, 2007; Eaglen *et al.*, 2012).

Table 22. Estimates (standard error) of (co)variance components and related genetic parameters from single-breed models using data from purebred data sets I and II

	BBB[a] purebreds		HOL[a] purebreds	
Parameter[c]	Model L1[b]	Model L2[b]	Model L1[b]	Model L2[b]
σ_h^2	0.06 (0.00)	0.06 (0.00)	0.04 (0.00)	0.04 (0.00)
σ_a^2	0.17 (0.01)	0.12 (0.01)	0.03 (0.00)	0.03 (0.00)
σ_m^2	0.05 (0.01)	0.02 (0.00)	0.01 (0.00)	0.01 (0.00)
σ_p^2	0.02 (0.00)	0.02 (0.00)	0.01 (0.00)	0.01 (0.00)
σ_e^2	0.24 (0.01)	0.26 (0.00)	0.25 (0.00)	0.25 (0.00)
$r_u(a,m)$	-0.46 (0.04)[d]	---	-0.29 (0.11)[d]	---
h_d^2	0.34 (0.02)[d]	0.25 (0.01)[d]	0.09 (0.01)[d]	0.07 (0.01)[d]
h_m^2	0.09 (0.01)[d]	0.05 (0.01)[d]	0.04 (0.01)[d]	0.03 (0.01)[d]
C_h	12%	12%	13%	13%
C_p	4%	4%	3%	3%
C_e	49%	54%	73%	74%
AIC[e]	-425324.10	108035.28	-218821.44	68273.02

[a] BBB = Belgian Blue breed, HOL= Holstein breed.

[b] Model L1 is the linear animal model with estimated genetic covariance between direct and maternal effects; Model L2 is the linear animal model with genetic covariance between direct and maternal effects constrained to zero described in Vanderick *et al.* (2014)

[c] The terms σ_h^2 is the herd × year of calving variance, σ_a^2 is the direct additive genetic variance, σ_m^2 is the maternal additive genetic variance, σ_p^2 is the permanent maternal environmental variance, σ_e^2 is the residual variance, $r_u(a,m)$ is the genetic correlation between direct and maternal effects, h_d^2 and h_m^2 are the direct and the maternal heritabilities, respectively. C_h, C_p and C_e are the herd × year of calving fraction, permanent maternal environmental fraction and residual fraction in the phenotypic variance, respectively.

[d] Approximated standard error computed using equations provided by Klei and Tsuruta (2008).

[e] AIC=Akaike's Information Criterion

A significant negative genetic correlation between direct and maternal additive effects was obtained in both breeds fitting Model L1 (Table 22). Negative correlations have been documented to exist between direct and maternal effects in beef and dairy cattle (*e.g.* Phocas and Laloë, 2004; Eaglen *et al.*, 2012). This negative direct-maternal genetic correlation suggests that dairy and beef farmers need to base selection decisions on both the direct and the maternal EBV of a sire for calving ease to avoid long-term negative consequences.

Direct and maternal additive variances as well as h^2 estimates showed obvious differences between breeds (BBB *vs* HOL). The BBB breed had greater direct and maternal additive variances and heritabilities than the HOL breed (Table 22). A trend for higher h^2 estimates of birth traits, such as calving ease, in beef breeds compared to dairy breeds was reported by Koots *et al.* (1994a). Moreover, a higher negative direct-maternal genetic correlation in BBB than in HOL was also found (-0.46 *vs* -0.29, respectively). As for the heritabilities, a trend for stronger negative direct-maternal genetic correlations in beef than in dairy breeds was reported by Koots *et al.* (1994b). Few differences in herd × year of calving and maternal permanent environmental variances were observed between breeds. These variances were slightly greater in BBB than in HOL (Table 22).

Model L1 had the best fit for both breeds, as shown by the smallest percentage of residual variance in the phenotypic variance (49% *vs* 54% in Model L2 for BBB, and 73% *vs* 74% for HOL). Moreover, according to the AIC values in Table 22, Model L1 presented the lower AIC value for BBB and HOL breeds, meaning that Model L1 was the best model to explain calving ease in both breeds. Therefore, for the multi-breed analysis, an estimated genetic correlation between direct and maternal additive genetic effects was considered in Model MBV and Model MBSM.

MULTI-BREED ANALYSIS

The estimates of (co)variance components and derived genetic parameters with their approximated SE are displayed in Table 23. Standard errors were relatively small for all estimates, except for estimates of genetic correlation between direct and maternal additive effects for HOL for both multi-breed models. Even, by using completely different approaches, estimates of (co)variance components and genetic parameters were similar for Model MBV and Model MBSM.

The estimates of segregation variances (Model MBSM) were 0.06 (±0.01) for direct effects and 0.01 (±0.00) for maternal effects; and high negative correlation between direct and maternal segregation effects was found (-0.56 ±0.03). The estimation of these variances and this correlation required informative records on the calves, for direct effects, and on the dams, for maternal effects, from advanced crosses (*i.e.* F2, third-cross (F3)...). In the present study, the number of these informative records did not allow an accurate estimation of the direct and maternal segregation variances and correlation since most of the crossbred animals were F1. Nevertheless, the related SE were low.

Because the different genetic parts were correlated in Model MBV, correlations among breed-specific additive genetic effects were estimated. Correlation between BBB and HOL was high for the direct genetic effects (0.90 ±0.42), whereas correlation for the maternal genetic effects was moderate (0.33 ±0.53) but with a large SE and hence, being unreliable. High genetic correlations should indicate that alleles inherited by purebreds express themselves similarly in crossbreds.

As in the single-breed analysis, evident genetic differences between BBB and HOL were observed within each multi-breed model. However, the magnitudes of direct and maternal h^2 estimates were quite different between single-breed and multi-breed analyses, especially for the HOL breed (Table 22 and Table 23). The h^2 estimates of the HOL population increased considerably by using multi-breed models in which purebred and crossbred data from both breeds were jointly analysed. In HOL, the direct h^2 increased 100% and 77% for Model MBV and Model MBSM, respectively, when compared to Model L1, whereas the maternal h^2

increased by 25% by using Model MBV or Model MBSM. Although these h² were higher than those used in the current Walloon genetic evaluation system (Vanderick *et al.*, 2013), these heritabilities were consistent with those reported in the literature for calving ease in Holstein dairy cattle, ranging from 0.03 to 0.17 for direct h² and from 0.02 to 0.12 for maternal h² (e.g. López de Maturana *et al.*, 2007; Eaglen *et al.*, 2012). To a lesser extent, an increase of direct h² for BBB animals was also observed by using multi-breed analyses (+32%). By contrast, a decrease of maternal h² (-14%) for BBB animals was observed between Model L1 and Model MBV and similar estimates of maternal h² were obtained comparing Model L1 with Model MBSM.

The direct-maternal genetic correlations increased in multi-breed models compared to single-breed models (+24% on average), except for the genetic direct and maternal correlation for HOL estimated *via* Model MBV (-22%). However, this latter correlation estimate was not reliable considering its large SE.

Based on the relative residual fraction of the phenotypic variance, even with small differences, we may infer that the best fit was achieved for Model MBV with 52 and 67% of relative residual variances for BBB and HOL animals, respectively. For Model MBSM these fractions were 53 and 69% respectively for BBB and HOL animals.

Table 23. Estimates (standard error) of (co)variance components and related genetic parameters from both multi-breed models using data from purebred herds and crossbred animals (data set III)

	Model MBV[a]		Model MBSM[b]	
Parameter[d]	BBB[c]	HOL[c]	BBB[c]	HOL[c]
σ_h^2	0.07 (0.02)	0.05(0.01)	0.07 (0.01)	0.05 (0.01)
σ_a^2	0.20 (0.09)	0.06 (0.03)	0.19 (0.10)	0.06 (0.03)
σ_m^2	0.04 (0.02)	0.02 (0.01)	0.04 (0.02)	0.02 (0.00)
σ_{sa}^2	---		0.06 (0.01)	
σ_{sm}^2	---		0.01 (0.00)	
σ_p^2	0.01 (0.01)	0.01 (0.00)	0.02 (0.01)	0.01 (0.00)
σ_e^2	0.24 (0.02)		0.23 (0.02)	
$r_u(a,m)$	-0.57 (0.09)[e]	-0.22 (0.23)[e]	-0.58 (0.08)[e]	-0.36 (0.17)[e]
$r_u(sa,sm)$	---		-0.56 (0.03)[e]	
h_d^2	0.45 (0.07)[e]	0.18 (0.05)[e]	0.45 (0.08)[e]	0.16 (0.04)[e]
h_m^2	0.08 (0.01)[e]	0.05 (0.01)[e]	0.09 (0.02)[e]	0.05 (0.01)[e]
C_h	15%	13%	15%	13%
C_p	3%	3%	3%	5%
C_e	52%	67%	53%	69%

[a] Model MBV is based on the random regression multi-breed model described in Vanderick *et al.* (2009)

[b] Model MBSM is based on the approximate split multi-breed model described in Strandén and Mäntysaari (2013)

[c] BBB = Belgian Blue breed, HOL= Holstein breed.

[d] The terms σ_h^2 is the herd × year of calving variance, σ_a^2 is the direct additive genetic variance, σ_m^2 is the maternal additive genetic variance, σ_{sa}^2 is the direct segregation variance, σ_{sm}^2 is the maternal segregation variance, σ_p^2 is the permanent maternal environmental variance, σ_e^2 is the residual variance, $r_u(a,m)$ is the genetic correlation between direct and maternal effects, $r_u(sa,sm)$ is the correlation between direct and maternal segregation effects, h_d^2 and h_m^2 are the direct and the maternal heritabilities, respectively. C_h, C_p and C_e are the herd × year of calving fraction, permanent maternal environmental fraction and residual fraction in the phenotypic variance, respectively.

[e] Approximated standard error computed using equations provided by Klei and Tsuruta (2008).

From the (co)variances estimated using multi-breed models, variances for calving ease were computed for BBB×HOL crossbred animals. Results are shown for F1 animals (BBB=50% and HOL=50%) and back-crosses (F2) between F1 and purebreds (BBB=25 or 75% and HOL=75 or 25%) in Table 24 and Table 25, respectively. The herd × year of calving and direct additive genetic variances for BBB × HOL crossbred calves were between those reported for purebreds in both multi-breed models (Table 23). Note that the estimates of direct additive genetic variances for BBB×HOL crossbred calves were larger in Model MBSM than Model MBV, especially for back-crosses. These greater estimates obtained for back-crosses might be explained by the inclusion of the direct segregation variances in Model MBSM, which is an extra source of genetic variability for those animals. The maternal additive genetic variances for BBB×HOL crossbred dams tended to follow a similar pattern (Table 25). By contrast, the maternal permanent environmental variances of crossbred dams were below those of the purebred dams in both multi-breed models as shown in Table 25.

Table 24. Variances of herd × year of calving and direct additive genetic effects estimated from both random regressions multi-breed models for purebred and crossbred calves

Breed composition of calf		Model MBV[a]		Model MBSM[b]	
BBB[a]	HOL[a]	Herd-period	Direct additive genetic	Herd-period	Direct additive genetic
0.00	1.00	0.05	0.06	0.05	0.06
0.25	0.75	0.05	0.08	0.04	0.12
0.50	0.50	0.05	0.11	0.05	0.12
0.75	0.25	0.06	0.15	0.07	0.19
1.00	0.00	0.07	0.20	0.07	0.19

[a] Model MBV is based on the random regression multi-breed model described in Vanderick *et al.* (2009)

[b] Model MBSM is based on the approximate split multi-breed model described in Strandén and Mäntysaari (2013)

Table 25. Variances of maternal permanent environmental and maternal additive genetic effects estimated from both random regressions multi-breed models for purebred and crossbred dams

Breed composition of dam		Model MBV[a]		Model MBSM[b]	
BBB[a]	HOL[a]	Maternal permanent environment	Maternal additive genetic	Maternal permanent environment	Maternal additive genetic
0.00	1.00	0.01	0.02	0.01	0.02
0.25	0.75	0.01	0.01	0.01	0.03
0.50	0.50	0.01	0.02	0.01	0.03
0.75	0.25	0.01	0.02	0.01	0.04
1.00	0.00	0.01	0.04	0.02	0.04

[a] Model MBV is based on the random regression multi-breed model described in Vanderick *et al.* (2009)

[b] Model MBSM is based on the approximate split multi-breed model described in Strandén and Mäntysaari (2013)

MODEL COMPARISON

The mean estimated reliabilities of EBV of the purebred sires and maternal grandsires obtained for each model are provided in Table 26. The two multi-breed models showed greater mean reliabilities than Model L1 for both breeds. Indeed, for Model MBV, the reliability gain was approximately 5 and 3 % for direct and maternal EBV in BBB, respectively, and around 25 and 15% in HOL. For Model MBSM, this reliability gain was approximately 5 and 6% for direct and maternal EBV in BBB, respectively, whereas were around 19 and 15% in HOL. By comparing multi-breed models, in terms of reliability, the Model MBV was similar to the Model MBSM for both breeds. It should also be highlighted that reliability depends upon genetic parameters, and because greater h^2 estimates were, in general, reported for multi-breed models (Table 23), the use of crossbred data jointly with purebred data in the genetic evaluation increased the reliability of the purebred EBV. This is in agreement with VanRaden *et al.* (2007) in a joint evaluation of purebreds and crossbreds in US dairy cattle and Lutaaya *et al.* (2002) in a joint evaluation of purebreds and crossbreds in swine. Both studies demonstrated that crossbreds are useful for improving the accuracy of genetic evaluation of their purebred relatives. Greater accuracy means greater genetic gain. Relative genetic gain from multi-breed models compared to Model L1 was estimated as the relative increase in the mean estimated accuracies (*i.e.* square root of mean estimated reliabilities) of EBV between Model L1 and multi-breed models. The relative advantages in genetic gain that could be expected by a breeder by using EBV from Model MBV instead of those from Model L1 ranged from 2% to 12% and from 3% to 9% by using EBV from Model MBSM instead of those from Model L1.

Table 26. Mean reliability of the estimated breeding values (EBV) of purebred sires for direct calving ease and of purebred maternal grandsires for maternal calving ease in both breeds

Breed[d]	EBV	N	Model L1[a]	Model MBV[b]	Model MBSM[c]
BBB	Direct	2395	0.55	0.58	0.58
	Maternal	2256	0.33	0.34	0.35
HOL	Direct	2696	0.31	0.39	0.37
	Maternal	2374	0.20	0.23	0.23

[a] Model L1 is the single-breed linear animal model described in Vanderick *et al.* (2014)

[b] Model MBV is based on the random regression multi-breed model described in Vanderick *et al.* (2009)

[c] Model MBSM is based on the approximate split multi-breed model described in Strandén and Mäntysaari (2013)

[d] BBB = Belgian Blue breed, HOL= Holstein breed.

The ranking of animals on the bases of their genetic merit to allow their selection is a major objective in breeding programs. Spearman's rank correlations between EBV obtained by Model L1, Model MBV and Model MBSM among purebred sires with progeny for direct calving ease, as well as among purebred maternal grandsires with progeny for maternal calving ease are presented in Table 27. The rank correlations were high (≥0.90) for all models in both breeds, indicating similar genetic rankings between all models, especially between Model MBV and Model MBSM, but also indicating that there will be no substantial losses regardless of which of the three models is used. A somewhat lower rank correlation among models was observed for HOL, especially between Model L1 and Model MBV. Furthermore, estimates of rank correlations were slightly smaller for maternal effects than for direct effects of calving ease in both breeds.

Table 27. Spearman's rank correlations of estimated breeding values (EBV) of purebred sires for direct calving ease and of purebred maternal grandsires for maternal calving ease in both breeds obtained from the different models[a]

Breed[d]	EBV	N	Model L1, Model MBV	Model L1, Model MBSM	Model MBV, Model MBSM
BBB	Direct	2395	0.97	0.97	0.99
	Maternal	2256	0.94	0.94	0.99
HOL	Direct	2597	0.96	0.95	0.99
	Maternal	2374	0.90	0.91	0.99

[a] Model L1 is the single-breed linear animal model described in Vanderick *et al.* (2014); Model MBV is based on the random regression multi-breed model described in Vanderick *et al.* (2009); Model MBSM is based on the approximate split multi-breed model described in Strandén and Mäntysaari (2013)

[b] BBB = Belgian Blue breed, HOL= Holstein breed.

Finally, all tested models were compared in terms of goodness of fit. To achieve this, the MSE and the Pearson's correlation estimates between observed and predicted calving ease scores of BBB and HOL purebred calves were calculated for each model. The results are presented in Table 28. Models with the smaller MSE had a better agreement between observed and fitted values, *i.e.* a better goodness of fit. The differences in MSE were very small between models, especially for BBB calves. A higher difference in MSE was observed for HOL purebred calves between Model L1 and Model MBV. The lowest MSE was obtained by Model MBSM for the BBB purebreds whereas the lowest MSE for HOL purebreds was obtained with Model MBV. Based on these MSE, both multi-breed models did perform slightly better than Model L1 in terms of goodness of fit; and therefore, accuracy of prediction was slightly better using Model MBV or Model MBSM than Model L1. Similar to MSE results, differences in Pearson's correlation between models were very small and again the multi-breed models outperformed single-breed models, especially for the HOL purebred calves (Table 28).

Table 28. Mean squared error (MSE) and Pearson's correlation estimates between observed and predicted calving ease scores from Model L1, Model MBV and Model MBSM for BBB and HOL purebred calves

	BBB[a] (N=55 319)		HOL[a] (N=40 535)	
	MSE	Correlation	MSE	Correlation
Model L1[b]	0.15	0.93	0.20	0.73
Model MBV[c]	0.14	0.93	0.17	0.79
Model MBSM[d]	0.14	0.93	0.18	0.78

[a] BBB = Belgian Blue breed, HOL= Holstein breed.

[b] Model L1 is the single-breed linear animal model described in Vanderick *et al.* (2014)

[c] Model MBV is based on the random regression multi-breed model described in Vanderick *et al.* (2009)

[d] Model MBSM is based on the approximate split multi-breed model described in Strandén and Mäntysaari (2013)

For a more detailed comparison of both multi-breed models, MSE and Pearson's correlation estimates were also computed for BBB × HOL crossbred calves. Crossbred calves were gathered in 5 groups according to their breed composition and results of MSE and Pearson's correlations for these 5 groups of crossbred calves are provided in Table 29.Although there were no large differences between the 2 multi-breed models, the MSE for the Model MBSM was smaller than for Model MBV in each of the 5 crossbred groups; therefore, Model MBSM had a slightly better goodness of fit for crossbred animals than Model MBV. The same statement can be made for the estimates of Pearson's correlation.

To sum up, Model MBV and Model MBSM worked almost equally well regarding comparison results for both purebred and crossbred animals. Therefore, both multi-breed models can be considered as being quasi-equivalent models and either of them can be used to perform the joint genetic evaluation for calving ease of BBB and HOL Walloon cattle.

Table 29. Mean squared error (MSE) and Pearson's correlation estimates between observed and predicted calving ease scores from Model MBV and Model MBSM for calves of data set III

Breed composition of calves			MSE		Correlation	
BBB [a]	HOL [a]	N	Model MBV[b]	Model MBSM[c]	Model MBV[b]	Model MBSM[c]
] 0% - 19%]	[80% - 95%[	564	0.29	0.27	0.93	0.93
[20% - 39%]	[60% - 79%]	1051	0.28	0.26	0.90	0.91
[40% - 59%]	[40% - 59%]	2504	0.24	0.22	0.85	0.86
[60% - 79%]	[20% - 39%]	435	0.45	0.40	0.90	0.92
[80% - 95%[	] 0% - 19%]	342	0.35	0.33	0.92	0.93

[a] BBB = Belgian Blue breed, HOL= Holstein breed.

[b] Model MBV is based on the random regression multi-breed model described in Vanderick *et al.* (2009)

[c] Model MBSM is based on the approximate split multi-breed model described in Strandén and Mäntysaari (2013)

CONCLUSIONS

According to estimates of genetic parameters, model fits and correlation between models, this study verified the feasibility of a joint genetic evaluation for calving ease trait using purebred and crossbred data from BBB and HOL Walloon cattle through multi-breed models. The adaptations of the two literature based multi-breed models (Model MBV and Model MBSM) were successfully demonstrated. In addition, it was shown that Model MBV and Model MBSM performed almost equally well for purebred and crossbred animal, *i.e.* quasi-equivalent. It was demonstrated that the use of BBB x HOL crossbred information had a positive influence on the estimation of genetic merit of BBB and HOL purebred animals, and accordingly on the relative genetic gain expected by breeders using EBV from both proposed multi-breed models. However, further studies are needed to assess the predictive ability of these multi-breed models.

These results showed the benefit of a joint genetic evaluation for calving ease of Walloon BBB and HOL cattle including crossbreds. This evaluation could be performed with either of the two proposed multi-breed models, since they were shown as quasi-equivalent models. But both models have advantages and disadvantages. The Model MBV allows describing the additive genetic effects of crossbreds as a weighted sum of breed-specific effects. We can therefore consider that the expectation of the additive effects to be correctly modelled. Nonetheless, Model MBV uses a strong simplification to define the additive variances whereas the Model MBSM defines in a theoretically more correct manner segregation effects and associated variances. Finally, the MBV model is computationally simpler and has the advantage that it could be easier to use in a genomic evaluation setting, even if Makgahlela *et al.* (2013) used the MBSM, but avoided difficulty in modelling segregation effects under genomic context.

ACKNOWLEDGMENTS

The Ministry of Agriculture of the Walloon Region of Belgium (Service Public de Wallonie, Direction générale opérationnelle “Agriculture, Ressources naturelles et Environnement) is acknowledged for its financial support. The authors gratefully acknowledge the technical support provided by the Walloon Breeding Association (Ciney, Belgium). Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI) funded by the National Fund for Scientific Research (F.R.S.-FNRS, Brussels, Belgium) under Grant No. 2.5020.11. The authors are grateful to G. R. Wiggans for his helpful suggestions and comments on the manuscript.

LITERATURE CITED

Abdel-Azim, G. A., and P. Berger. 1999. Properties of threshold model predictions. J. Anim. Sci. 77:582–590.

Barrier, A. C., M. J. Haskell, S. Birch, A. Bagnall, D. J. Bell, J. Dickinson, A. I. Macrae, and C. M. Dwyer. 2013. The impact of dystocia on dairy calf health, welfare, performance and survival. Vet. J. 195:86–90.

Bennett, G., and K. Gregory. 2001. Genetic (co) variances for calving difficulty score in composite and parental populations of beef cattle: I. Calving difficulty score, birth weight, weaning weight, and postweaning gain. J. Anim. Sci. 79:45–51.

Berry, D. P., and R. D. Evans. 2014. Genetics of reproductive performance in seasonal calving beef cows and its association with performance traits. J. Anim. Sci. 92:1412–1422.

Eaglen, S. A. E., M. P. Coffey, J. A. Woolliams, and E. Wall. 2012. Evaluating alternate models to estimate genetic parameters of calving traits in United Kingdom Holstein-Friesian dairy cattle. Genet. Sel. Evol. 44:1–13.

Fouz, R., F. Gandoy, M. L. Sanjuán, E. Yus, and F. J. Diéguez. 2013. The use of crossbreeding with beef bulls in dairy herds: effects on calving difficulty and gestation length. animal 7:211–215.

García-Cortés, L. A., and M. Á. Toro. 2006. Multibreed analysis by splitting the breeding values. Genet. Sel. Evol. 38:601–615.

Hagger, C., and A. Hofer. 1990. Genetic analyses of calving traits in the Swiss Black and White, Braunvieh and Simmental breeds by REML and MAPP procedures. Livest. Prod. Sci. 24:93–107.

Hansen, L. 2006. Monitoring the worldwide genetic supply for dairy cattle with emphasis on managing crossbreeding and inbreeding. In: Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Minas Gerais, Brazil, 13-18 August, 2006. Instituto Prociência. p. 1–1.

Heins, B., L. Hansen, and A. Seykora. 2006. Calving difficulty and stillbirths of pure Holsteins versus crossbreds of Holstein with Normande, Montbeliarde, and Scandinavian Red. J. Dairy Sci. 89:2805–2810.

Klei, B., and S. Tsuruta. 2008. Approximate variance for heritability estimates. Available from: http://nce.ads.uga.edu/~shogo/html/research/AI_SE.pdf

Kolkman, I., S. De Vliegher, G. Hoflack, M. Van Aert, J. Laureyns, D. Lips, A. de Kruif, and G. Opsomer. 2007. Protocol of the Caesarean Section as Performed in Daily Bovine Practice in Belgium. Reprod. Domest. Anim. 42:583–589.

Koots, K. R., J. P. Gibson, C. Smith, and J. W. Wilton. 1994a. Analyses of published genetic parameter estimates for beef production traits. I. Heritability. Anim. Breed. 62:309–338. (Abstr.).

Koots, K. R., J. P. Gibson, and J. W. Wilton. 1994b. Analyses of published genetic parameter estimates for beef production traits. II. Phenotypic and genetic correlations. Anim. Breed. 62:825–853 (Abstr.).

Lo, L. L., R. L. Fernando, and M. Grossman. 1993. Covariance between relatives in multibreed populations: additive model. Theor. Appl. Genet. 87:423–430.

López de Maturana, E. L., A. Legarra, L. Varona, and E. Ugarte. 2007. Analysis of Fertility and Dystocia in Holsteins Using Recursive Models to Handle Censored and Categorical Data. J. Dairy Sci. 90:2012–2024.

Lutaaya, E., I. Misztal, J. W. Mabry, T. Short, H. H. Timm, and R. Holzbauer. 2002. Joint evaluation of purebreds and crossbreds in swine. J. Anim. Sci. 80:2263–2266.

Makgahlela, M. L., E. A. Mäntysaari, I. Strandén, M. Koivula, U. S. Nielsen, M. J. Sillanpää, and J. Juga. 2013. Across breed multi-trait random regression genomic predictions in the Nordic Red dairy cattle. J. Anim. Breed. Genet. 130:10–19.

Mayeres, P., C. Michaux, M. Szydlowski, C. Bertozzi, and N. Gengler. nd. Application of mixed major gene-polygenic inheritance model for birth traits in dual purpose Belgian Blue cattle.

Mee, J. F. 2008. Prevalence and risk factors for dystocia in dairy cattle: A review. Vet. J. 176:93–101.

Misztal, I., S. Tsuruta, D. Lourenço, I. Aguilar, A. Legarra, and Z. Vitezica. 2014. Manual for BLUPF90 family of programs. Athens Univ. Ga.

Phocas, F., and D. Laloë. 2003. Evaluation models and genetic parameters for calving difficulty in beef cattle. J. Anim. Sci. 81:933–938.

Phocas, F., and D. Laloë. 2004. Genetic parameters for birth and weaning traits in French specialized beef cattle breeds. Livest. Prod. Sci. 89:121–128.

Sørensen, M. K., E. Norberg, J. Pedersen, and L. G. Christensen. 2008. Invited Review: Crossbreeding in Dairy Cattle: A Danish Perspective. J. Dairy Sci. 91:4116–4128.

Strandén, I., and E. A. Mäntysaari. 2013. Use of random regression model as an alternative for multibreed relationship matrix. J. Anim. Breed. Genet. 130:4–9.

Vanderick, S., B. L. Harris, J. E. Pryce, and N. Gengler. 2009. Estimation of test-day model (co)variance components across breeds using New Zealand dairy cattle data. J. Dairy Sci. 92:1240–1252.

Vanderick, S., T. Troch, A. Gillon, G. Glorieux, P. Faux, and N. Gengler. 2013. Genetic evaluation of calving ease for Walloon Holstein dairy cattle. Interbull Bull.:32–37.

Vanderick, S., T. Troch, A. Gillon, G. Glorieux, and N. Gengler. 2014. Genetic parameters for direct and maternal calving ease in Walloon dairy cattle based on linear and threshold models. J. Anim. Breed. Genet. 131:513–521.

VanRaden, P. M., and A. H. Sanders. 2003. Economic merit of crossbred and purebred US dairy cattle. J. Dairy Sci. 86:1036–1044.

VanRaden, P. M., M. E. Tooker, J. B. Cole, G. R. Wiggans, and J. H. Megonigal. 2007. Genetic Evaluations for Mixed-Breed Populations. J. Dairy Sci. 90:2434–2441.

CHAPTER V. GENERAL DISCUSSION, CONCLUSIONS AND PERSPECTIVES

CHAPTER V. GENERAL DISCUSSION, CONCLUSIONS AND PERSPECTIVES

OUTLINE

Chapters II to IV presented genetic evaluation systems for the genetic improvement of maternally influenced traits and genetic evaluation models for the genetic improvement of crossbred and purebred animals from multi-breed populations. The objectives of this chapter are to compile results obtained throughout this work and to discuss them in a wider perspective. The following aspects relating to the accuracy and efficiency of genetic selection based on estimated breeding values are addressed. Firstly, the quantity and quality of phenotypic data used in any genetic evaluation are examined. Secondly, the use of adequate statistical genetic models to get accurate estimations of genetic parameters and estimated breeding values is discussed. Finally, main conclusions obtained from previous chapters as well as future perspectives of research are summarized.

INTRODUCTION

Animal breeding programs are set up to genetically improve livestock populations to enhance competitiveness and sustainability of livestock production. Genetic selection based on estimated breeding values (EBV) is key in genetic improvement programs, since it provides a way of comparing animals to select the genetically superior indiviuals that will be used as parents of the next generation. Estimated breeding values are random effect solutions of mixed model equations (Henderson, 1973) that are usually based on the knowledge of phenotypic data and pedigree information. In the last ten years, the inclusion of molecular information (*e.g.* SNP markers), in addition to phenotypes and pedigree, has become an important component in the prediction of genetic merit.

The mixed model equations can be set up in different ways depending on the evaluated trait, the population analysed, the final goal and other considerations, and in summary, the accuracy of genetic selection depends not only on quantity and quality of the available phenotypic data, but also on the suitability of the statistical genetic evaluation (mixed) model used for the estimation of genetic parameters and EBV.

PHENOTYPIC DATA

QUANTITY OF DATA

Phenotypic and pedigree data sets are the foundation of genetic evaluation systems. Phenotypic data may be recorded by the animals' performance recording systems (*e.g.* milk recording organisations), progeny-test stations, animal breeding organisations (*e.g.* Herd-Book), livestock farmers (on-farm recording), scientific research centres, laboratories (*e.g.* milk testing labs), veterinarians, slaughterhouses and so on.

The quantity of recorded data relies on many factors such as the amount of labour and its cost, drive of farmers and other stakeholders to record more data, the availability of computing resources, etc. Depending on the trait to be evaluated, the data quantity can be huge (*e.g.* several tens of gigabytes of data for milk yield or spectral data…) or limited to a few hundreds of phenotypes, especially if the recording is difficult and/or expensive. For instance, direct recording of traits such as dry matter intake or methane emission in a large number of animals is currently hard to achieve, which limits the efficiency of genetic selection for such traits.

The number of phenotypes depends heavily upon the incentive of the breeders and producers to record the traits. For example, although recording of calving ease scores is not mandatory, the quantity of scores recorded by breeders has considerably increased since the genetic evaluation system of calving ease for Walloon Holstein dairy cattle was implemented. From December 2012 (first official genetic evaluation for calving ease) to December 2016 (most recent genetic evaluation), an increase of 94% was observed (from 85 118 to 165 085 scores) whereas only 85 118 scores relevant for genetic evaluation were recorded between 2000 and 2012 (Chapter III.2). This significant increase can, at least partially, be explained by the fact that the delivery of EBV to Walloon dairy farmers during the 4 last years. The usefulness of these EBV as decision support tools for selection and mating, which in turn has made the farmers more willing to supply extra phenotypic data to increase the reliability of the EBV.

In addition, the amount of available data is increasing due to developments of precision livestock farming, defined as the management of livestock production using principles and technologies from process engineering. The fast advances in computer resources, information technology as well as the implementation of electronic animal identification and the development of the sensor-based data, have greatly influenced data recording in livestock populations. The use of advanced sensors closely associated to animals facilitates therefore the recording of data at low cost without disturbing animals and can help more effectively farmers in managing their herds (Gengler *et al.*, 2013). On dairy farms, sensors measuring milk conductivity or pedometer measuring mobility are often given as examples. Likewise, fine milk composition, and in particular fatty acid profiles, can be measured using mid-infrared (MIR) spectroscopy, which is an example of high-throughput data with respect to both the number of measurements and the number of variables. Thus, precision livestock farming has the potential to increase dramatically the quantity of phenotypic data available.

QUALITY OF DATA

As previously mentioned, the accuracy of estimated genetic parameters and EBV rely on the quality of phenotypic data and on correct pedigree recording. Moreover, high quality data are crucial for making efficient use of genomic data in applications such as linkage or association mapping and genomic selection. A prerequisite to accurate genetic/genomic analysis is data consistency to ensure data quality.

Quality checks are performed on data to identify inconsistent or inaccurate data that should not be used in a genetic analysis. The aim of the data cleaning process is to ensure that the analysis is as accurate as possible while the data still constitute a representative sample of the population. Quality checks were carried out in Chapters III and IV. For example, records considered as outliers or individuals with missing information related to any factor used in the genetic analysis model (*e.g.* animal/herd identification, birth/calving date) were removed. Likewise, birth dates of animals and parentage were checked to ensure that the age of an animal at recording and its parents' ages at its birth were consistent. Data recording may be optional in some cases (*e.g.* calving ease scores in Wallonia), which introduces uncertainty in data quality due to subjectivity or errors from breeders. To avoid these issues, some additional checks must be applied to the data. For example, herds with a standard deviation for calving ease scores ≤ 0.05 were excluded to avoid herds where breeders score all calving in the same category (Chapters III.1, III.2 and IV.2). Similarly, to avoid breeder recording errors, survival limits based on the mean of lamb survival to weaning for a flock and birth year were applied on the lamb survival data (Chapter III.3).

Furthermore, information from other recorded traits can be used to check the relevance/consistency of data for a trait under evaluation, thus improving the data quality for this trait (*e.g.* Kadarmideen and Coffey, 2001). In Chapter III.3, the presence/absence of lamb weight measurements recorded at the strategically important time periods combined with information coming from lamb birth fate codes was used to assess whether a lamb had survived until weaning. Based on this combination of data, new decision rules for data inclusion and exclusion in the genetic evaluation system of lamb survival to weaning were defined and tested. Results showed that these new rules allowed to enhance data quality and to recover data that were previously considered as missing when using the previous decision rules.

As was the case with the quantity of phenotypic data, the incentive of farmers and other stakeholders involved in recording is essential to ensure good quality of phenotypic data. In addition, precision farming can also improve data quality, especially through the use of electronic identification based on radio frequency. The information recorded and stored on the reader is more accurate than using pen and paper, with fewer opportunities for errors to occur. Once recorded and stored, the information can be easily transferred for further use. Thus, the use of sensors or other electronic devices (*e.g.* automated milking systems, automated weighing crates, feeding machines) combined with electronic identification enables the acquisition of the good quality phenotypic data (*e.g.* Boichard and Brochard, 2012).

To sum up, the quality of data used in a genetic evaluation system can be guaranteed and improved through the implementation of quality checks, using other traits as information sources, properly incentivising farmers and using precision farming tools.

STATISTICAL GENETIC EVALUATION MODEL

Another important step in genetic evaluation systems is to choose an appropriate statistical model to estimate genetic parameters and genetic merit (*i.e.* EBV) accurately.

The choice of a genetic evaluation model may be based on whether trait under evaluation is expressed on a continuous or on a discrete scale, as well as being informed by the underlying processes affecting the expression of the trait, especially in the presence of multiple environmental (such as permanent environment, maternal or not) or genetic (such as animal and maternal genetics or non-additivity) factors. Furthermore, the model needs to be adapted to the particular structure of the animal population analysed (*e.g.* multi-breed population, admixed population).

MODELS FOR TRAITS EXPRESSED ON DISCRETE SCALES (CATEGORICAL TRAITS)

Many traits of importance in livestock production are expressed on discrete scales that are categorical (*e.g.* calving ease, disease susceptibility, survival). Therefore, genetic evaluation models accounting for the categorical nature of the trait is needed.

Theoretically, non-linear mixed models, such as generalized linear mixed models (GLMMs) and threshold mixed models, allow better estimation of genetic parameters and genetic merit predictions of categorical traits (Gianola, 1982). Researchers in general expected that non-linear mixed models would lead to increased response from selection, because they described more accurately the structure of the data (Abdel-Azim and Berger, 1999). However, although some simulation studies (Hoeschele, 1988; Abdel-Azim and Berger, 1999) have confirmed the statistical superiority of non-linear mixed models for analysing discrete data, several studies using phenotypic data from sheep, beef, and dairy cattle demonstrated no real significant advantage for the analysis of categorical traits (Matos *et al.*, 1997; Phocas and Laloë, 2003; Vazquez *et al.*, 2012).

Our results agree with the literature. In Chapter III.1, calving ease scores from Walloon Holstein dairy cattle were analysed using linear and threshold animal mixed models. Models were assessed and compared in terms of goodness of fit and predictive ability. Although threshold models had better results in terms of goodness of fit, no clear advantage of the threshold mixed models over the linear mixed models was observed in terms of predictive ability, especially when there were a larger number of progeny per sire (Table 4). Furthermore, Spearman's rank correlations of 0.97 were found between sires EBV indicating that sire rankings were similar and that fitting linear models would not result in a substantial loss of accuracy. Likewise, when revising the genetic evaluation model of the lamb survival to weaning trait for the New Zealand sheep industry (Chapter III.3), the benefits of using use a generalized linear mixed model with a logit link function rather a linear mixed model were investigated. Results obtained from the cross-validation study showed that, in terms of predictive ability, the generalized linear mixed model (MSE from 0.1283 to 0.1293) was slightly less accurate than the linear mixed model (MSE from 0.1275 to 0.1284) on lamb survival data. These results could, at least partially, be explained, by the hypothesis that in many practical situations, linear models are naturally very robust to non-normality. Similarly, non-linear models also often rely on specific assumptions that are not necessarily adequately met.

Our results combined to previous studies in literature suggested little incentive for the use of non-linear over linear mixed models of both categorical traits studied, especially because computational requirements are more complex by using non-linear mixed models (Chapter II). However, the use of non-linear models might be worthwhile for some types of categorical traits, as those showing unordered or few categories (e.g., binomial), very uneven distributions of records inside categories or non-normal probability distributions (e.g. Poisson).

MODELS FOR MATERNALLY INFLUENCED TRAITS

Several traits of interest in animal production systems (*e.g.* calving ease, birth weight, survival to weaning) are maternally influenced (*e.g.* Dematawena and Berger, 1997; Roughsedge *et al.*, 2001). These traits are influenced not only by the genotype or the environment of the individual itself (*i.e.* direct effects), but also by either environment and/or the genotype of its mother (*i.e.* maternal effects). Genetic evaluations of maternally influenced traits require models accounting for all those effects to deliver unbiased and accurate EBV. Access to accurate maternal genetic evaluation systems could indeed help breeders and producers to select superior animals for breeding, thereby improving the maternal efficiency of the livestock. Estimation of direct genetic as well as maternal genetic and environmental effects is possible by fitting a maternal animal model (Chapter II).

Through Chapters III.1 to III.3, different maternal animal models were investigated and assessed for two maternally influenced traits: calving ease in Walloon Holstein population (Chapters III.1 and III.2) and lamb survival to weaning in New Zealand sheep population (Chapter III.3). In Chapters III.1 and III.2, Models L1 and L2 were linear models and Models T1 and T2 were threshold models. Models L1 and T1 were fitted with an estimated correlation between direct and maternal additive effects whereas Models L2 and T2 were fitted with a correlation between these effects constrained to zero. In Chapter III.3, models A1, A2, A3 and B were linear models whereas models Log A and LogB were logistic models using model A3 and model B, respectively. Models A1 and A2 differed in the way they handled the variable "lamb day of birth", model B was the same as A3 but added a maternal environmental random effect.

An important issue in the implementation of a genetic evaluation system for a maternally influenced trait is to determine the relationship between direct and maternal genetic effects. This relationship affects the estimation of genetic parameters as well as the response to selection of both components. Most of genetic correlation estimates between direct and maternal effects tend to be negative (*e.g.* Bennett and Gregory, 2001; Phocas and Laloë, 2003; Eaglen and Bijma, 2009; Everett-Hincks *et al.*, 2014), although positive correlations have also reported in some studies (*e.g.* Matos *et al.*, 2000; Steinbock *et al.*, 2003). These negative correlations might be the product of both an existing genetic antagonism and a sire by year interaction (Robinson, 1996), and can be also influenced by data structure (Maniatis and Pollott, 2003). In our studies, positive and negative genetic correlation estimates between direct and maternal effects were obtained. For both studied maternally influenced traits, the sign of genetic correlation estimates differed between the linear and non-linear models (Tables 2, 11 and 12). However, almost all estimates were non-significant, except for estimates from linear models in Chapter III.3 (Table 11). Consequently, no genetic correlation between direct and maternal effects is fitted in the current Walloon genetic evaluation system for calving ease for Holstein dairy cattle (Chapter III.2). It should be noted that, during the single-breed analysis addressed in Chapter IV.2, a significant negative correlation estimate between direct and maternal additive genetic effects were obtained for calving ease for the same population with Model L1 (Table 22). This might be due to a better data structure and a greater amount of score phenotypes, as well as due to different definition of purebred Holstein animals between our studies (at least 75% of Holstein genes in Chapter III.1 *vs* at least 95% in Chapter IV.2). In the near future, we could consider a revision of the current genetic evaluation model and therefore to revisit the issue of the relationship between those effects for this trait. If a negative genetic correlation between direct and maternal effects for calving ease is confirmed, this will indicate that Walloon dairy farmers will need to base selection decisions on both direct and maternal EBV of an animal, to avoid long-term negative consequences. Currently, both direct and maternal EBV are provided separately to the breeders, which may cause some breeders to select their animals based solely on the direct (or conversely on the maternal) EBV. Calculating an aggregated EBV including direct and maternal values as suggested by Eaglen *et al.* (2012) could therefore be an effective way to encourage breeders to emphasise both effects, especially if strong negative correlations are present.

The necessity to include maternal effects, especially a maternal additive genetic effect, in genetic evaluation models can be seen by comparing estimated direct heritability (h^2) and total h^2, which was calculated as the ratio of total heritable variance (*i.e.* genetic variance that is available for response to selection) over phenotypic variance (Eaglen and Bijma, 2009). For lamb survival to weaning (Chapter III.3), this comparison showed that the presence of maternal effects led to an increase in estimated total genetic variance of approximately 75-100% in the case of models A1 to A3 whereas a decrease of 13% was observed with model B. This decrease for model B could be explained by the decrease in maternal additive genetic variance after maternal environmental effects (or litter effects) were added to the model, and when moderate negative genetic correlation between direct and maternal effects (-0.438) were estimated. The estimated total genetic variance was increased by approximately 255% and 175% for logistic models (Log A and Log B), respectively. On the other hand, for calving ease (Chapter III.1), the estimated total h^2 ranged from 0.102 to 0.151 for all tested models whereas the direct h² ranged from 0.074 to 0.121. It indicates that the addition of maternal effects increased heritable variance by approximately 25% for threshold models and by approximately 35% for linear models.

Furthermore, the importance of accounting for maternal environmental effects to prevent the overestimation of maternal h² and the subsequent overestimation of total h² was reported in several studies (*e.g.* Arango *et al.*, 2005; Everett-Hincks *et al.*, 2014). Our results agree with their findings. In Chapter III.3, higher estimates of maternal h² for lamb survival to weaning trait were observed in maternal models ignoring the maternal environmental effects (*i.e.* models A1 to A3 and Log A) compared to maternal models (*i.e.* models B and LogB) fitting those effects. Similar results were obtained during preliminary analysis of calving ease (Chapters III.1). In addition, the largest part of maternal variability was due to environmental effects rather than additive effects for both traits (Table 2 and Tables 11-12), reinforcing the importance of accounting for maternal environmental effects.

Finally, genetic models developed and assessed in Chapter III to evaluate calving ease and lamb survival to weaning were single-trait models. However, several authors showed that multiple-trait models may be better than single-trait models (*e.g.* Eaglen *et al.*, 2012; Everett-Hincks *et al.*, 2014). They suggested that using a multiple-trait model that incorporates a highly heritable and correlated indicator trait would improve the estimation of genetic parameters and would lead to a more optimal analysis. In particular, maternal variance for lowly heritable traits, such as calving ease or lamb survival, benefited from including a correlated trait (*e.g.* gestation length, body condition score (BCS), type traits, birth weight) in the genetic model (Bastin *et al.*, 2010; Eaglen *et al.*, 2013; Ahlberg *et al.*, 2016). Thus, it might be worthwhile to see whether there is real benefit in replacing the current single-trait model used to evaluate calving ease in the Walloon dairy cattle by a multiple-trait model using correlated traits, such as gestation length or BCS.

MODELS FOR MULTI-BREED (CROSSBRED) POPULATION

Crossbreeding has been shown to provide a simple method to increase the health and productivity of many plants and animals. It is widely used in the pig, poultry and beef cattle populations while, in general, it is not yet the case in dairy cattle. An exception can be made for New Zealand and for tropical climates where crossbreeding is wide spread in dairy cattle populations. However, for several years, problems related to functional traits (*e.g.* lower fertility, greater susceptibility to mastitis, leg problems and ketosis), coupled with increased inbreeding on an international scale, have resulted in tremendous interest in crossbreeding among commercial dairy producers (Sørensen *et al.*, 2008). Dairy producers indeed use crossbreeding to exploit heterosis for functional traits and breed complementarity for other traits. Therefore, dairy cattle populations that include crossbred animals can contribute to genetic progress. This can be achieved, for instance, by including crossbred bulls used in New Zealand in some populations (Chapter IV.1).

In general, genetic evaluation systems applied to livestock production, especially dairy cattle, only compare animals within breed. It means that only purebred data are used whereas data recorded on crossbred animals are neglected, even when it is available. We stated in chapter II that using purebred selection is not appropriate in the case of crossbreeding systems. There are indeed genetic and environmental differences between purebred and crossbred animals, so much so that purebred performance can be a poor predictor of crossbred performance (Ibánẽz-Escriche *et al.*, 2011). Hence, to perform combined crossbred-purebred selection, multi-breed genetic evaluation systems combining purebred and crossbred data have been suggested by numerous authors (*e.g.* Pollak and Quaas, 1998; Lutaaya *et al.*, 2002; García-Cortés and Toro, 2006; VanRaden *et al.*, 2007). Thus, a direct comparison of animals with various breed composition from multi-breed or admixed populations can be made, enabling genetic improvement of these populations. Moreover, with the introduction of genomic selection in animal breeding programs, many investigations have been carried out recently to develop multi-breed models allowing the use of molecular information, such as SNP markers (Toosi *et al.*, 2010; Makgahlela *et al.*, 2013; Christensen *et al.*, 2014; Lund *et al.*, 2014). Relatively sophisticated models need to be developed, for instance, to separate markers into 2 classes:

those with effects consistent across breeds, and those with breed specific effects or to allow marker effects to be partitioned into across and within breed components (Lund *et al.*, 2014).

In Chapter II, it was stated that in the case of a multi-breed population it is more appropriate to use breed specific additive genetic (co)variances. Results from single-breed analyses (Chapter IV) indicated genetic differences between the breeds studied (Holstein *vs* Jersey and Holstein *vs* Belgian Blue, respectively), which prompted the development in Chapter IV.1 and IV.2 of multi-breed models with heterogeneous genetic (co)variances across breeds to conserve the genetic features of each breed. The particularity of these multi-breed models is that they use breed proportions to estimate random regression coefficients. Random regressions are useful because they can accommodate changes in the (co)variance structure of phenotypic data according to the breed composition of an animal. Therefore, the EBV of an animal was decomposed into several correlated genetic parts depending on their breed of origin. This decomposition of the genetic (co)variance matrix by breed of origin resulted in a much simpler formulation for (co)variance components estimation, which is easy to implement using estimation techniques available in general purpose software.

The multi-breed model proposed in Chapter IV.2 (called hereafter Model MBV for convenience) was compared with the approximate multi-breed random regression model presented by Strandén and Mäntysaari (2013) (called hereafter Model MBSM), which also used breed-specific random regressions for the additive genetic random effects but based on the square root of breed proportions. The main difference between Models MBV and MBSM was how they handled the relationships between breeds in the additive genetic (co)variance structure. Contrary to Model MBV, covariances between breed-specific additive genetic effects were not accounted for by Model MBSM, most likely because they were already included in the segregation (co)variances structure. The segregation variance results from differences in allelic frequencies between parental pure breeds, and is derived as the difference in additive variances between breed groups (Chapter II). Results showed the quasi-equivalence of both Models MBV and MBSM in terms of goodness of fit as well as in terms of genetic rankings (Spearman’s rank correlation of 0.99), especially for F1 crossbreds when the segregation effect is null. Likewise, the same kind of comparison was undertaken between the multi-breed models (*i.e.* Model MBV and Model MBSM) and the single-breed model used for the single-breed analysis (Chapter IV.2). Results indicated similar genetic rankings between single-breed and Model MBV or Model MBSM, but also indicated there would be no substantial losses in accuracy incurred by using either of the multi-breed models.

According to several studies, one of the main advantages of the multi-breed models, which make use of purebred and crossbred information, was the increased reliabilities of the purebred EBV compared to strictly single-breed models, due to the use of all available progeny information (Lutaaya *et al.*, 2001; VanRaden *et al.*, 2007; Ibánẽz-Escriche *et al.*, 2011). The same conclusions were reached in Chapter IV.2 between single-breed models using purebred data alone and multi-breed models (Model MBV or Model MBSM) combining purebred and crossbred data. Results showed that the use of data from crossbred progeny in addition to purebred data resulted in reliability gains ranging from 3% up to 25% depending on breed and genetic additive effect, and confirming the importance of using data from crossbred progeny in addition to purebred data. Thus, since response to selection is proportional to the prediction accuracy of the genetic merit, combining purebred and crossbred information may be beneficial for genetic evaluation of purebred animals.

Finally, results from Chapter IV indicated that breed-specific EBV for a trait measured repeatedly over time (*i.e.* milk yield in first lactation) as well as for a maternally influenced trait (*i.e.* calving ease) could be estimated for each purebred and crossbred animals of a multi-breed population using the multi-breed random regression models proposed. These breed-specific EBV could be helpful for breeders and producers in their selection and mating decisions because they could select animals based on the mating strategy that they would like to apply in their herds.

Although Model MBV and Model MBSM performed almost equally well for purebred and crossbred animals (*i.e.* quasi-equivalent models), they both have advantages and disadvantages. The Model MBV allows the additive genetic effects of crossbreds to be described as a weighted sum of breed-specific effects. We can therefore consider the expectation of the additive effects to be correctly modelled. However, Model MBV uses a strong simplification to define the additive variances whereas the Model MBSM's definition of segregation effects and associated variances is theoretically more correct. Finally, the MBV model is computationally simpler since the estimation of segregation variance/effects is not required, so that it has the advantage that it could be easier to use in a genomic evaluation framework. Indeed, the principal difference between single-breed and multiple-breed genomic prediction is in the relationship matrix used to relate SNP effects to phenotypes. The genomic relationship matrix needs to properly account for the breed of origin of the alleles in the population as well as the allele frequency differences between breeds, and this could easily be managed by Model MBV. However, most studies have found little benefit in using models that included breed-specific genetic effects over those which ignored these effects (*e.g.* Makgahlela *et al.*, 2013; Thomasen *et al.*, 2013; Lourenco *et al.*, 2016). But most studies in dairy cattle have focused on the use of multi-breed animals to predict the genetic merit of purebred animals, rather than using all available information to the predict genetic merit of composite animals. It would therefore be interesting to modify Model MBV to extend the utility of genomic selection to those crossbred populations.

GENERAL CONCLUSIONS

The main conclusions drawn from this thesis are that:

- For both categorical traits studied (*i.e.* calving ease and lamb survival to weaning), non-linear mixed models did not offer clear advantages over linear mixed models. No substantial losses in accuracy will result from fitting linear rather than threshold mixed models for these categorical traits.

- For both maternally influenced traits analysed (*i.e.* calving ease and lamb survival to weaning), the largest part of maternal variability was due to environmental effects. Therefore, maternal environmental effects had to be considered in the genetic models used to evaluate calving ease and lamb survival to weaning to get accurate estimation of additive genetic effects. Indeed, it was shown that genetic models omitting maternal environmental effects (temporary or permanent) resulted in an overestimation of maternal additive genetic variance and, thus in an overestimation of maternal h^2.

- For both maternally influenced traits analysed, positive and negative genetic correlations between direct and maternal effects were estimated through linear and threshold maternal models. Almost all estimates of genetic correlations were not significant, with two exceptions. Estimates from linear maternal models for lamb survival to weaning showed moderate and unfavourable correlations. Moreover, a significant negative correlation was estimated for calving ease during the development of a joint genetic evaluation of this trait for Holstein and Belgian Blue Walloon cattle.

- Based on the linear maternal animal models developed in this thesis, useful and accurate EBV for direct and maternal additive genetic effects can be estimated, allowing breeders and producers to improve the genetic merit of their livestock herds efficiently.

- Genetic correlations across breeds computed from the multi-breed random regression models showed additive genetic differences between breeds as well as differences in the additive effects transmitted to purebred and crossbred offspring.

- Breed-specific EBV could be estimated by the multi-breed random regression models for purebred and crossbred animals from multi-breed populations and, therefore, help breeders and producers to select animals according to their mating strategy.

- Genetic evaluation using multi-breed models combining data recorded on purebred and crossbred animals led to an increase in the reliability of EBV for purebred animals.

PERSPECTIVES

This thesis contributed to the development of genetic evaluation systems currently used in routine genetic evaluations in Walloon Region of Belgium and in New Zealand. It also opened further directions of research:

- to revise the current genetic evaluation model of calving ease for Walloon dairy cattle as the amount of recorded data has increased significantly since the implementation of the system, leading to a need to re-examine if the correlation between direct and maternal genetic effects can still be considered equal to zero of if it needs to be re-estimated.

- to investigate the usefulness of a multiple-trait maternal animal model using correlated traits rather than the current single-trait maternal animal model used to evaluate calving ease for the Walloon dairy cattle. A first step would be to examine genetic relationship between calving ease and other traits, such as gestation length or BCS, to determine which trait could be a good indicator trait to enhance the accuracy of genetic selection and therefore increase the genetic gains in calving ease.

- to extend the proposed multi-breed random regression model to more breeds.

- to include dominance effects in the proposed multi-breed random regression model.

- to modify the multi-breed random regression model for dense SNP markers to allow genomic selection in multi-breed populations.

IMPLICATIONS

Researches undertaken during this thesis have led to the development and the implementation of a genetic evaluation system for calving ease in the Walloon Region of Belgium. Since April 2013, EBV for direct and maternal effects of calving ease from Holstein animals have been provided to Walloon breeders and producers and can be used for their breeding decisions. Moreover, the development of a genetic evaluation system for calving ease has allowed the Walloon Region of Belgium to participate in the international MACE evaluation for calving traits performed by Interbull. Finally, since April 2015, direct and maternal calving ease EBV have been integrated in the Walloon global selection objective (V€G).

Likewise, researches carried out during my scientific stay at the AgResearch Limited, Invermay Agricultural Centre (Mosgiel, New Zealand) have contributed to the development of the current genetic evaluation system for lamb survival to weaning in New Zealand.

LITERATURE CITED

Abdel-Azim, G. A., and P. Berger. 1999. Properties of threshold model predictions. J. Anim. Sci. 77:582–590.

Ahlberg, C. M., L. A. Kuehn, R. M. Thallman, S. D. Kachman, W. M. Snelling, and M. L. Spangler. 2016. Breed effects and genetic parameter estimates for calving difficulty and birth weight in a multibreed population. J. Anim. Sci. 94:1857–1864.

Arango, J., I. Misztal, S. Tsuruta, M. Culbertson, and W. Herring. 2005. Estimation of variance components including competitive effects of Large White growing gilts. J. Anim. Sci. 83:1241–1246.

Bastin, C., S. Loker, N. Gengler, A. Sewalem, and F. Miglior. 2010. Short communication: Genetic relationship between calving traits and body condition score before and after calving in Canadian Ayrshire second-parity cows. J. Dairy Sci. 93:4398–4403.

Bennett, G., and K. Gregory. 2001. Genetic (co) variances for calving difficulty score in composite and parental populations of beef cattle: I. Calving difficulty score, birth weight, weaning weight, and postweaning gain. J. Anim. Sci. 79:45–51.

Boichard, D., and M. Brochard. 2012. New phenotypes for new breeding goals in dairy cattle. Animal 6:544–550.

Christensen, O. F., P. Madsen, B. Nielsen, and G. Su. 2014. Genomic evaluation of both purebred and crossbred performances. Genet. Sel. Evol. 46:23.

Dematawena, C. M. B., and P. J. Berger. 1997. Effect of dystocia on yield, fertility, and cow losses and an economic evaluation of dystocia scores for Holsteins. J. Dairy Sci. 80:754–761.

Eaglen, S. A. E., and P. Bijma. 2009. Genetic parameters of direct and maternal effects for calving ease in Dutch Holstein-Friesian cattle. J. Dairy Sci. 92:2229–2237.

Eaglen, S. A. E., M. P. Coffey, J. A. Woolliams, and E. Wall. 2012. Evaluating alternate models to estimate genetic parameters of calving traits in United Kingdom Holstein-Friesian dairy cattle. Genet. Sel. Evol. 44:1–13.

Eaglen, S. A. E., M. P. Coffey, J. A. Woolliams, and E. Wall. 2013. Direct and maternal genetic relationships between calving ease, gestation length, milk production, fertility, type, and lifespan of Holstein-Friesian primiparous cows. J. Dairy Sci. 96:4015–4025.

Everett-Hincks, J. M., H. C. Mathias-Davis, G. J. Greer, B. A. Auvray, and K. G. Dodds. 2014. Genetic parameters for lamb birth weight, survival and death risk traits. J. Anim. Sci. 92:2885–2895.

García-Cortés, L. A., and M. Á. Toro. 2006. Multibreed analysis by splitting the breeding values. Genet. Sel. Evol. 38:601–615.

Gengler, N., D. Berry, and C. Bastin. 2013. Use of automated systems for recording of direct and indirect data with special emphasis on the use of MIR milk spectra (OptiMIR project). ICAR Tech. Ser. No 17:55.

Gianola, D. 1982. Theory and analysis of threshold characters. J. Anim. Sci. 54:1079–1096.

Henderson, C. R. 1973. SIRE EVALUATION AND GENETIC TRENDS. J. Anim. Sci. 1973. Available from: http://dx.doi.org//1973.1973Symposium10x

Hoeschele, I. 1988. Comparison of "Maximum A-Posteriori Estimation" and "Quasi Best Linear Unbiased Prediction" with threshold characters. J. Anim. Breed. Genet. 105:337–361.

Ibán̋ez-Escriche, N., J. Reixach, N. Lleonart, and J. L. Noguera. 2011. Genetic evaluation combining purebred and crossbred data in a pig breeding scheme. J. Anim. Sci. 89:3881–3889.

Kadarmideen, H. N., and M. P. Coffey. 2001. Quality and validation of insemination data for national genetic evaluations for dairy cow fertility in the United Kingdom. Interbull Bull.:133.

Lourenco, D. A. L., S. Tsuruta, B. O. Fragomeni, C. Y. Chen, W. O. Herring, and I. Misztal. 2016. Crossbreed evaluations in single-step genomic best linear unbiased predictor using adjusted realized relationship matrices. J. Anim. Sci. 94:909–919.

Lund, M. S., G. Su, L. Janss, B. Guldbrandtsen, and R. F. Brøndum. 2014. Genomic evaluation of cattle in a multi-breed context. Livest. Sci. 166:101–110.

Lutaaya, E., I. Misztal, J. W. Mabry, T. Short, H. H. Timm, and R. Holzbauer. 2001. Genetic parameter estimates from joint evaluation of purebreds and crossbreds in swine using the crossbred model. J. Anim. Sci. 79:3002–3007.

Lutaaya, E., I. Misztal, J. W. Mabry, T. Short, H. H. Timm, and R. Holzbauer. 2002. Joint evaluation of purebreds and crossbreds in swine. J. Anim. Sci. 80:2263–2266.

Makgahlela, M. L., E. A. Mäntysaari, I. Strandén, M. Koivula, U. S. Nielsen, M. J. Sillanpää, and J. Juga. 2013. Across breed multi-trait random regression genomic predictions in the Nordic Red dairy cattle. J. Anim. Breed. Genet. 130:10–19.

Maniatis, N., and G. E. Pollott. 2003. The impact of data structure on genetic (co) variance components of early growth in sheep, estimated using an animal model with maternal effects. J. Anim. Sci. 81:101–108.

Matos, C. A. P., D. L. Thomas, L. D. Young, and D. Gianola. 2000. Genetic analyses of lamb survival in Rambouillet and Finnsheep flocks by linear and threshold models. J. Anim. Sci. 71:227–234.

Matos, C. A., D. L. Thomas, D. Gianola, M. Perez-Enciso, and L. D. Young. 1997. Genetic analysis of discrete reproductive traits in sheep using linear and nonlinear models: II. Goodness of fit and predictive ability. J. Anim. Sci. 75:88–94.

Phocas, F., and D. Laloë. 2003. Evaluation models and genetic parameters for calving difficulty in beef cattle. J. Anim. Sci. 81:933–938.

Pollak, E. J., and R. L. Quaas. 1998. Multibreed genetic evaluations of beef cattle. In: Proc 6th World Congr Genet Appl Livest Prod. Vol. 23. p. 81–88. Available from: http://www.bifconference.com/bif2005/pdfs/Pollak.pdf

Robinson, D. L. 1996. Models which might explain negative correlations between direct and maternal genetic effects. Livest. Prod. Sci. 45:111–122.

Roughsedge, T., R. Thompson, B. Villanueva, and G. Simm. 2001. Synthesis of direct and maternal genetic components of economically important traits from beef breed-cross evaluations. J. Anim. Sci. 79:2307–2319.

Sørensen, M. K., E. Norberg, J. Pedersen, and L. G. Christensen. 2008. Invited Review: Crossbreeding in Dairy Cattle: A Danish Perspective. J. Dairy Sci. 91:4116–4128.

Steinbock, L., A. Näsholm, B. Berglund, K. Johansson, and J. Philipsson. 2003. Genetic effects on stillbirth and calving difficulty in Swedish Holsteins at first and second calving. J. Dairy Sci. 86:2228–2235.

Strandén, I., and E. A. Mäntysaari. 2013. Use of random regression model as an alternative for multibreed relationship matrix. J. Anim. Breed. Genet. 130:4–9.

Thomasen, J. R., A. C. Sørensen, G. Su, P. Madsen, M. S. Lund, and B. Guldbrandtsen. 2013. The admixed population structure in Danish Jersey dairy cattle challenges accurate genomic predictions. J. Anim. Sci. 91:3105–3112.

Toosi, A., R. L. Fernando, and J. C. M. Dekkers. 2010. Genomic selection in admixed and crossbred populations. J. Anim. Sci. 88:32–46.

VanRaden, P. M., M. E. Tooker, J. B. Cole, G. R. Wiggans, and J. H. Megonigal. 2007. Genetic Evaluations for Mixed-Breed Populations. J. Dairy Sci. 90:2434–2441.

Vazquez, A., M. Perez-Cabal, B. Heringstad, M. Rodrigues-Motta, G. Rosa, D. Gianola, and K. Weigel. 2012. Predictive ability of alternative models for genetic analysis of clinical mastitis. J. Anim. Breed. Genet. 129:120–128.

LIST OF TABLES

LIST OF FIGURES