ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

AgroParisTech
INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE L'ENVIRONNEMENT
PARIS INSTITUTE OF TECHNOLOGY FOR LIFE, FOOD AND ENVIRONMENTAL SCIENCES

Ecole Doctorale ABIES
Agriculture
Agriculture
Alimentation
Food
Biologie
Biology
Environnement
Environment
Santé
Health

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

L'Institut des Sciences et Industries du Vivant et de l'Environnement

(AgroParisTech)

Spécialité : Génétique animale

présentée et soutenue publiquement par

Clotilde PATRY

le 9 décembre 2011

Impacts of genomic selection on classical genetic evaluations

Les impacts de la sélection génomique sur les évaluations génétiques classiques

Directeur de thèse : **Vincent Ducrocq**

*Représentant CIFRE UNCEIA : **Laurent Journaux***

Jury

M. Etienne VERRIER, Professeur, AgroParisTech, France	Président
M. Erling STRANDBERG, Professeur et chef de département, SLU, Suède	Rapporteur
M. Zengting LIU, Chef de projets Recherche et Développement, VIT, Allemagne	Rapporteur
M. Laurent JOURNAUX, Responsable du service génétique, UNCEIA, France	Examinateur
M. Vincent DUCROCQ, Directeur de Recherche, INRA GABI, France	Examinateur

AgroParisTech
INRA GABI
Domaine de Vilvert, 78350 Jouy-en-Josas, France

Abstract

In 2011, genomic evaluations have been implemented in 16 countries and at least 7 dairy cattle breeds. Selection decisions are no longer based on breeding values estimated after progeny testing but on those obtained after genomic predictions. In less than 4 years, the use of these early and accurate genomically enhanced breeding values (GEBV) was integrated in breeding schemes. This new strategy is called genomic selection.

Applied to mixed linear models, Best Linear Unbiased Prediction (BLUP) is the method used for national genetic evaluations and its extension, the Multiple Across Country Evaluation (MACE) is performed at the international level. With genomic selection, the selection process is no longer fully described in mixed model equations: only the selected candidates have actual progeny records and are included in the analysis. This violates the implicit assumption of random Mendelian sampling contribution used in BLUP applied to an animal model. Breeding value estimates are feared to be biased and less accurate. Because of incomplete and incorrect national information, MACE solutions would also be affected.

The adoption of genomic tools is fast and worldwide. It is therefore important to consider the consequences of a genomic selection step in breeding schemes on the current genetic evaluations. This is relevant at the national and international levels.

Based on real data in the Holstein population, bias was assessed by simulation. It was first measured in a French national evaluation and then on three country scales after international evaluations. Different levels of selection intensity and of genetic correlations between country scales were tested.

These simulations showed evidence of significant biases in classical evaluations. The breeding values of the genomically selected young sires were clearly underestimated at the national level and estimates were less accurate. These cohorts were also the most penalized in international rankings. They were even more affected when national breeding values were not only incomplete but also incorrect. In fact, bias could propagate to relatives and foreign populations on the different country scales. It is not only necessary but urgent to account for the genomic selection step in national evaluations.

The main requirement is to include all genotyped candidates in national evaluations, i.e., the selected and the culled ones based on their GEBV. One approach was tested to do so: the GEBV of all candidates were converted into pseudo-performances and associated with an appropriate weight derived from the genomic reliability. It was implemented under simplistic assumptions and gave satisfactory results as it totally eliminated the bias at the national level. However, it is not so straightforward to include genomic information in classical evaluations: some genomic information might be redundant.

Other approaches are proposed to first make GEBV reliabilities more realistic. Besides multi-step approaches for genetic evaluations, a single step procedure including all available information in the same analysis would be more satisfactory. Genomic information is more properly distributed to the whole population. Computational strategies based on an iterative procedure were suggested to implement it. This could be the most optimal solution to quickly prevent from bias and at the same time to enhance the accuracy of national evaluations. It still needs to be tested. It is also necessary to develop tests to check whether national evaluations are unbiased after genomic selection before including them in international evaluations.

A main consequence of bias corrected breeding values is that all breeding values will include some genomic information in a near future. This is not yet accepted in international evaluations: the latter need to be improved and adapted to deal with GEBV.

Key-words

Genomic Selection - Genetic Evaluations - BLUP - MACE - Bias - Dairy Cattle - Interbull

This research was conducted with the financing from **Agence Nationale de la Recherche** (ANR,project AMASGEN) and **APIS-GENE**, at:

Institut National de la Recherche Agronomique (INRA)

Génétique Animale et Biologie Intégrative (GABI)
Equipe G2B, génétique et génomique des bovins laitiers et allaitants
Domaine de Vilvert
78350 Jouy-en-Josas
France

With the support of :

Union Nationale des Coopératives d'Elevage et d'Insémination Animale (« CIFRE » contract n°0679/2008, **Association Nationale de la Recherche de la Technologie** fundings)

UNCEIA

Service Génétique
Maison Nationale des Eleveurs
149, rue de Bercy
75595 Paris Cedex 12
France

In collaboration with:

Interbull centre
Dept. Animal Breeding and Genetics
Swedish University of Agricultural Sciences (SLU)
Box 7023
75007 Uppsala
Sweden

A mon père,
Pour sa passion, son combat et ses paradoxes

"[The captain] had bought a large map representing the sea,
Without the least vestige of land:
And the crew were much pleased when they found it to be
A map they could all understand."

Lewis Carroll – The hunting of the Snark

.

Abbreviations

AI	Artificial or animal insemination
BLUP	Best Linear Unbiased Prediction
DGV	Direct Genomic Value
DRP	De-Regressed Proof
DYD	Daughter Yield Deviation
EBV	Estimated Breeding Value
(g)EDC	(Genomic) Effective Daughter Contributions
EN	Elastic-Net
GEBV	Genomically Enhanced Breeding Value
G-BLUP	BLUP using a Genomic relationship matrix
HMME	Henderson Mixed Model Equations
Interbull	International Bull Evaluation Service
MACE	Multiple Across Country Evaluation
MS	Mendelian Sampling
MAS	Marker Assisted Selection
MSE	Mean Squared Errror
PA	Parent Average
PEV	Predictor Error Variance
PT	Preferential Treatment
QTL	Quantitative Trait Loci
R^2	Reliability
RP	Reference Population
SNP	Single Nucleotide Polymorphism
TBV	True Breeding Value

CONTENTS

Abstract 2

Key-words 3

Abbreviations 9

CHAPTER 1 - General Introduction 14

1.1. Background: the emergence of genomics in dairy cattle breeding 14

1-1.1. An additional type of information for breeding value estimation 14

1-1.2. The potential of genomic information: an increased genetic gain 14

1-1.3. A fast and wide development of genomic evaluations 14

1-1.4. Changes in genetic evaluations, breeding strategies, and worldwide organisation 15

1-1.5. Semantic issues 15

1.2. Context of the thesis: the AMASGEN research project 16

1.3. Aim of the thesis 16

1.4. Outlines of the thesis 17

CHAPTER 2 - Genomic selection in dairy cattle: a turning point in breeding value estimation and in breeding strategies 20

2.1. Improvement of genetic evaluations using additional information 20

2-1.1. The use of an increasing amount of information 20

2-1.2. The use of a new type of information, the genomic one 25

2-1.3. Combination of genomic and polygenic information: how to get GEBV? 27

2.2. New breeding strategies to use genomic information 32

2-2.1. Principle and logistics of a regular progeny test scheme 32

2-2.2. Moving away from progeny testing schemes 32

2-2.3. Optimization of breeding schemes based on genomic information 33

2-2.4. Changes in semen distribution 35

CHAPTER 3 - The reasons for bias in classical evaluations after a genomic selection step ... 36

3.1. Best and unbiased genetic predictions 36

3-1.1. Definitions of bias and precision of an estimator 36

3-1.2. Statistical properties of BLUP solutions 37

3-1.3. Assumptions for optimal properties of BLUP 38

3.2. Cases of biases in genetic evaluations..... 39

3-2.1. Biases due to selection in classical evaluations..... 39

3-2.2. Other sources of bias in classical evaluations 39

3-2.3. Mechanisms of bias in genetic evaluations..... 40

3-2.4. The use of molecular information: new types of selection bias 41

3.3. Bias in classical evaluations after a genomic selection step 42

3-3.1. Genomic selection: the combined effects of "selective phenotyping" and "sequential selection" 42

3-3.2. Violation of the mixed model assumptions: biased BLUP solutions 42

3-3.3. Bias in MACE solutions 42

3-3.4. A worry shared by the international community 43

CHAPTER 4 - Bias assessment in national evaluations after a genomic selection step..... 44

4.1. Assessment of bias in breeding value estimates..... 44

4-1.1. Measures of bias and accuracy 44

4-1.2. Factors influencing the direction and magnitude of bias..... 45

4.2. Our simulation study 46

4-2.1. Outlines of the implemented method 46

4-2.2. Summary of the simulation results..... 47

Article I 48

CHAPTER 5 - Accounting for a genomic selection step in national evaluations..... 60

5.1. Necessity to account for bias..... 60

5-1.1. Long terms problems due to bias in genetic evaluations 60

5-1.2. Bias due to genomic selection: a large-scale and urgent problem..... 60

5.2. Approaches to account for selection 62

5-2.1. Usual approaches to account for bias in BLUP evaluations 62

5-2.2. Alternative approaches to account for genomic selection..... 64

5.3. Our implemented strategy 65

5-3.1. Choices for de-regression method and gEDC computation..... 65

5-3.2. Alternative scenarios..... 65

5-3.3. Efficiency of the method in reducing bias magnitude 66

5-3.4. Main conclusions of the simulation study *66*

Article II 68

CHAPTER 6 - Bias propagation in international evaluations due to genomic selection 80

6.1. The roles of Interbull: changes due to genomic selection 80

6-1.1. The roles of Interbull *80*

6-1.2. New Interbull's roles in the genomic era *81*

6.2. The importance of international evaluations in the genomic era 82

6.3. Risks of biases in international evaluations 83

6-3.1. Various strategies between countries to implement genomic selection *83*

6-3.2. Two main sources of bias in international evaluations *83*

6.4. Bias assessment in international evaluations 83

6-4.1. General strategy *83*

6-4.2. Summary of the simulation results *84*

Article III 86

CHAPTER 7 - General Discussion 106

7.1. Limits and contributions of the study 106

7-1.1. Limits of the simulation strategies and analyses *106*

7-1.2. A better understanding of the bias mechanisms after simulations of genomic selection *107*

7.2. Consequences on classical and genomic evaluations 109

7-2.1. The extent of bias in classical evaluations *109*

7-2.2. Widespread problems from biased classical evaluations *111*

7.3. Techniques and methods to avoid bias propagation 114

7-3.1. Methods to prevent bias in national genetic evaluations *114*

7-3.2. Possible implementations of these approaches in routine evaluations *120*

7-3.3. Techniques to avoid bias propagation in international genetic evaluations *124*

CHAPTER 8 - General Conclusion 128

8.1. Impact of genomic selection on genetic evaluations 128

8-1.1. An established risk for genetic evaluations *128*

8-1.2. An urgent need to avoid such bias *128*

8-1.3. A relatively easy approach to implement at the national level *129*

8.2. A broader scope of the questions related to bias 129

8-2.1. Integration of genomic information in classical evaluations: objectives beyond the bias issue *130*

8-2.2. Perspectives: genomic information propagated in all estimations of breeding values *130*

8-2.3. Unsolved issues at the international level: incompatibilities with MACE *131*

8.3. Perspectives for international collaborations in the dairy cattle breeding organization . 132

8-3.1. Economical and technological context: a high competitiveness *132*

8-3.2. Opportunities and fears generated by genomic selection *132*

8-3.3. The on-going changes in the organization of the world dairy cattle breeding *133*

8-3.4. New market strategies and players *133*

REFERENCES 136

Table of illustrations 142

Publications 144

Peer-reviewed publications *144*

Other publications *144*

Co-authored publications *144*

Résumé 146

Mots-clés 150

CHAPTER 1 - General Introduction

1.1. Background: the emergence of genomics in dairy cattle breeding

1-1.1. AN ADDITIONAL TYPE OF INFORMATION FOR BREEDING VALUE ESTIMATION

After the first sequencing of the whole bovine genome in 2006, hundreds of thousands of DNA markers in the form of Single Nucleotide Polymorphisms (SNP) were discovered throughout the genome. It dramatically increased the pool of usable molecular information to enhance the quality of genetic evaluations in dairy cattle. At the same time, the dramatic reduction in genotyping costs (costs have been divided by 1 million in ten years!) made this technology available for routine use in dairy breeding schemes.

1-1.2. THE POTENTIAL OF GENOMIC INFORMATION: AN INCREASED GENETIC GAIN

Genetic progress in dairy cattle depends on the accuracy of genetic evaluations, the intensity of selection and the generation interval. The potential for molecular information to improve the rate of genetic gain has been known for decades (Smith, 1967, Soller and Beckmann, 1983). Marker-assisted selection has improved somewhat the efficiency of breeding programs (Dekkers, 2004) but its implementation has been very limited (Boichard *et al.*, 2010). The accessible marker genotypes information for males and females now provide early and accurate information at reasonable costs. Breeding values can now be estimated using the genome-wide dense information and then used as early selection tools. This is called genomic selection. Its implementation offers great potentials for increasing and managing a genetic gain.

1-1.3. A FAST AND WIDE DEVELOPMENT OF GENOMIC EVALUATIONS

In 2006, Powell and Norman reported that "*we may be past the time when there will be revolutionary changes in genetic evaluation techniques. […] Changes will likely be incremental, each addressing some shortcoming of the prior system*." Just one year later, Goddard and Hayes (2007) reviewed the potential for genomic selection and reported that "*Widespread use of DNA markers will have a major impact on the structure of the breeding programs and a significant impact on production systems more generally*." In 2008, genomic evaluations were run in 8 countries (Canada, Denmark, France, the Netherlands, New-Zealand, the USA, and Sweden) for the Holstein breed. In 2009, a first international workshop about "*Genomic Information in Genetic Evaluations*" was organized by Interbull (International Bull Evaluation Service) in Uppsala (Sweden). In 2010, genomic evaluations

were performed for 13 populations including 5 dairy breeds and 10 countries (Nilforooshan *et al.*, 2010). Now, 16 countries can be identified as having implemented genomic evaluations: to the first 8 ones, we can add Germany in 2009, Finland, Austria, Ireland, Switzerland and Poland in 2010 and at least Italy and UK in 2011. Japan also developed genomic evaluations at a research stage. Moreover, genomic evaluations are not only performed for the Holstein breed but also for the Brown Swiss, the Fleckvieh, the Jersey, the Mondbéliarde, the Normande and the Red Nordic breeds. Some authors (Hayes *et al.*, 2009) reported that "*genomic selection is revolutionizing dairy cattle breeding*".

1-1.4. CHANGES IN GENETIC EVALUATIONS, BREEDING STRATEGIES, AND WORLDWIDE ORGANISATION

Genomic technologies are changing genetic evaluation techniques. New statistical and computational developments have emerged to analyze genomic information and enhance the quality of genetic predictions.

To fully benefit from the opportunities generated by genomic selection, the design of breeding strategies must be adapted to take advantage of the early and accurate genomically enhanced breeding values. New breeding strategies have been developed and implemented in less than 4 years moving away from progeny testing. Selection objectives have been or are being reviewed to put more emphasis on functional traits with low heritability but with better perspectives for improvement due to genomic selection (Ducrocq, 2010). Expectations of breeding companies and farmers also include a wide access to cheaper and efficient genotyping technologies as well as the development of innovative herd management tools.

At the same time, the dairy cattle breeding world is moving toward an unbalanced market with competitive advantages for companies implementing genomic selection. In just a few years, the development of genomic evaluations and the implementation of genomic selection have led to a high competitiveness worldwide between not only breeding companies but also between research/computing centers for genetic predictions. In such a context, the necessity to keep on computing international genetic evaluations (as routinely delivered by the Interbull centre) was first questioned before to be reinforced because of their strategic role in multinational reference populations created for accurate genomic predictions.

1-1.5. SEMANTIC ISSUES

Genomics is a developing field and a large diversity of expressions can be found in the literature. The term "*Genomic Selection*" was first introduced by Haley and Visscher at the World Congress of Genetic Applied to Livestock Production (6th WCGALP) in 1998 and refers either to a field of research (Goddard and Hayes, 2007) or an approach to estimate breeding values (Meuwissen *et al.*, 2001) or a tool for selection decisions (Hayes *et al.*, 2009). In the present manuscript, "*genomic selection*" is explicitly differentiated from "*genomic evaluation*". The procedure that delivers estimated breeding values based on molecular

information will be called "*genomic evaluation*" whereas "*genomic selection*" will refer to the selection decisions based on these genomic breeding values.

Another source of misunderstanding is the name given to evaluations based on the use of phenotypes and pedigree only. Scientists call these either "traditional", "conventional" or "classical" evaluations, but also "BLUP evaluations", "polygenic evaluations" or simply "genetic evaluations". The first three alternatives are not precise and tend to be pejorative. The other ones are not valid strictly speaking: **BLUP** method for "Best Linear Unbiased Predictions" is the most widely implemented method to estimate breeding values based on phenotypic and pedigree information. However, genomic predictions may also rely on BLUP-derived methods. They may also include polygenic information, i.e., information on phenotypes and pedigree, and genomic predictions are now one particular step of the genetic evaluation process.

None of these denominations are strictly correct but for the sake of clarity, "classical evaluation" is the term chosen in this manuscript. Each type of evaluations, the genomic and the classical ones, will be considered as a step of the process for delivering an estimation of breeding values and the entire process will be called "genetic evaluation".

1.2. Context of the thesis: the AMASGEN research project

In France, a research project called AMASGEN was launched in 2009 by INRA (the French research institute for agriculture) with the collaboration of professionals from the federation of breeding and AI (artificial or animal insemination) cooperatives (Union Nationale des Coopératives d'Elevage et d'Insémination animale, UNCEIA). The main aim of this project was to develop a method to combine genomic information from genotyped animals with the information from phenotypes and pedigree for a fast and large implementation of the genomic selection in the French dairy cattle breeding schemes. The fifth and last work package of this project was dedicated to the aim of this study.

1.3. Aim of the thesis

With the emergence of the genomic era, it has become relevant to consider the consequences of a genomic selection step in breeding schemes on the current genetic evaluations at the national and international levels. This motivated the present study focusing on four main questions:

- Does genomic selection impact the classical predictions of animal breeding values computed using linear mixed models and how?
- How can we measure potential biases, i.e., systematic over- or underestimations of breeding values, in classical genetic evaluations?
- How large is this bias and is it necessary to develop approaches to reduce it?

- Which methods can therefore be proposed to address this issue?

1.4. Outlines of the thesis

Chapter 2 of this thesis aims at highlighting the recent changes in prediction methods and on design of breeding schemes due to the availability of molecular information. This chapter first presents how in the past the gain in amount of information for the estimation of breeding values motivated the adaptation of statistical models and methods. The respective BLUP and MACE (Multiple Across Country Evaluation) methods for the national and international evaluations are described. The principles of genomic evaluations are described as well as the current approaches suggested to combine genomic and classical evaluations. The estimation of breeding values has been strongly based on the implementation of progeny testing programs. This chapter also depicts their usual organization and how breeding programs are currently reconsidered to take advantage of genomic information.

Chapter 3 explains why a genomic selection step could be an issue in genetic evaluations. The statistical definition of bias is given and the properties of BLUP are described. Biases in BLUP solutions have been a recurrent problem in the past. They have often been caused by missing information after selection. With the access to molecular information, new types of bias have appeared and are reviewed. This chapter also explains why not only BLUP but also MACE solutions might be biased after implementation of a genomic selection step.

Chapter 4 focuses on the bias assessment in national evaluations after genomic selection. Experiences from the past can help in identifying the factors influencing bias direction and magnitude and in suggesting methods to measure such a bias. Based on real and simulated data, one strategy is proposed and applied. Methods and results for assessing bias in BLUP solutions are presented in a first article: **Patry, C. and V. Ducrocq. 2011. Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. J Dairy Sci 94:1011-1020** (article I).

Chapter 5 explains why it is necessary (and how) to account for genomic selection in national evaluations. The potential scope of bias due to genomic selection might become large enough to make necessary finding ways to avoid it. BLUP adjustments from past experiences and recent propositions to account for genomic selection are reviewed. A strategy is defined, relying on the simulation framework developed to assess bias in article I. Methods and results of this approach are presented in a second article: **Patry, C. and V. Ducrocq. 2011. Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. Genet Sel Evol 43:30** (article II).

In chapter 6, the importance of avoiding the propagation of bias at the international level is examined. International evaluations are delivered by Interbull. The roles of this organisation and how they are changing are described. International genetic evaluations are important for the accuracy of genomic predictions. So, it is relevant to prevent international

breeding values from being biased. Two issues are identified: completeness and correctness of the data sets provided by participating countries to Interbull. Based on real and simulated data, bias due to genomic selection and its spread are assessed in international evaluations. Methods and results are presented in a third article: **Patry, C., H. Jorjani and V. Ducrocq. Implementation of genomic selection at national level: impact of pre-selected and biased national BLUP evaluations on international genetic evaluations, submitted to J Dairy Sci on September, 27th , 2011** (article III).

A general discussion of the methods and results follows in chapter 7. The first part is dedicated to the limits and contributions of the methods implemented to measures bias in national and international evaluations after genomic selection. In the second part, the impact of such a bias in genetic evaluations is evaluated from the statistician's but also from the breeder's point of view. The relevance of the implemented method to account for genomic selection in breeding schemes at the national level is discussed as well as ways to improve it based on alternative approaches. Some propositions are addressed to avoid the propagation of bias in international genetic evaluations

Chapter 8 concludes on the opportunities and risks generated by a bias in the genetic evaluations due to genomic selection. The consequences on the world dairy breeding organization are especially examined.

CHAPTER 2 - Genomic selection in dairy cattle: a turning point in breeding value estimation and in breeding strategies

The objective of this chapter is to understand how approaches dedicated to breeding value estimation and how dairy cattle breeding strategies are currently evolving to benefit from a new type of information, i.e., the genomic information coming from dense molecular markers, in the case of dairy cattle.

2.1. Improvement of genetic evaluations using additional information

2-1.1. THE USE OF AN INCREASING AMOUNT OF INFORMATION

This section presents from a historical perspective the incremental evolutions of statistical models and methods for estimation of breeding values. Many changes occurred during the late 20th century but the following description only refers to the elements participating in the analysis of an increasing number of records in order to enhance the quality of genetic evaluations. Models, variance components and methods are systematically described using matrix notations as in Mrode (2005) who published an extensive overview of linear models for the prediction of animal breeding values.

- **Statistical basis for breeding value estimation**

Prediction of breeding values is a key element to implement breeding schemes. The quality of genetic evaluations usually depends on the availability of trait records. Data for genetic evaluations in dairy cattle are commonly collected at national level.

Variations among records or phenotypic observations are explained by environmental and genetic factors. Mixed linear models are chosen to analyze records: environmental effects are considered as fixed whereas the genetic effects and the residuals are considered as random. The statistical model for genetic predictions is, in matrix notation:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \quad [1]$$

where $\mathbf{y}$ is the vector of records, $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices relating animals to effects, $\mathbf{b}$ is a vector of fixed environmental effects, $\mathbf{u}$ and $\mathbf{e}$ are vectors of genetic and residual random effects.

The BLUP method, for Best Linear Unbiased Prediction, was developed by Henderson (1963, 1973) and relies on the simultaneous estimation of fixed and random effects, accounting for genetic relationships among the animals to evaluate. This method has been widely

implemented because of its optimal properties: BLUP solutions are known to be unbiased and to be the most accurate predictors among the linear functions of the data. BLUP was first implemented under a sire model, i.e., **s** replaces **u** in [1]. BLUP equations are also called Henderson's Mixed Model Equations (**HMME**) and are usually written as:

$$\begin{bmatrix} \mathbf{X'R^{-1}X} & \mathbf{X'R^{-1}Z} \\ \mathbf{Z'R^{-1}X} & \mathbf{Z'R^{-1}Z} + \alpha_s \mathbf{A^{-1}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{s}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'R^{-1}y} \\ \mathbf{Z'R^{-1}y} \end{bmatrix} \quad [2]$$

R and **G** are the residual and sire genetic variance-covariance matrices, **A** is the numerator relationship matrix for sires such as the variance of sire effect is: $\text{var}(\mathbf{s}) = \mathbf{A}\sigma_s^2$ with σ_s being the sire standard deviation of the trait. α_s is the variance ratio between residual and sire variances: $\alpha_s = \frac{\sigma_e^2}{\sigma_s^2} = \frac{4 - h^2}{h^2}$ with h^2 the trait heritability.

In the history of livestock improvement, BLUP-based methods to estimate breeding values have been continuously improved to include more information and increase accuracy of the estimation under computing constraints (Powell and Norman, 2006). One major objective is to deliver early estimated breeding values (**EBV**) with sufficient accuracy, so that selection decisions can be made as soon as possible.

- **From sire to animal models**

The additive genetic value *a* of an animal *i* depends on its parent breeding values (a_s and a_d) and on its Mendelian sampling contribution (φ), being the result of recombination and segregation of sire and dam alleles during gamete formation:

$$a_i = \frac{1}{2} a_s + \frac{1}{2} a_d + \varphi_i \quad [3]$$

Initially, only sires were evaluated with a sire model and the genetic merit of their mate, the dam (*d*) was not accounted for. This led to bias in the predicted breeding values when genetic trends started to generate heterogeneity in the genetic level of dams. One of the first advances has been to consider all relationships so that all animals, males and females, are simultaneously evaluated. Animal model has often replaced sire model. Since 1992, animal model has been considered as the standard method for national evaluation systems (Interbull, 1992):

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e} \quad [4]$$

a is the vector of random animal effects with $\text{var}(\mathbf{a}) = \mathbf{A}\sigma_a^2$, σ_a being the genetic standard deviation of the trait. The corresponding variance ratio is $\alpha = \frac{\sigma_e^2}{\sigma_a^2} = \frac{1 - h^2}{h^2}$. Note that $\sigma_a^2 = 4\sigma_s^2$

.

y is the vector of actual or pre-corrected records. For cows, pre-corrected records may consist in "Yield Deviations" which are weighted averages of cow's performances corrected for all

effects except genetic merit and residuals. For sires, it is common to use Daughter Yield Deviations (**DYD**) which are weighted average of daughter records corrected for all fixed effects and the breeding values of their dam (Van Raden and Wiggans, 1991). To define the daughter contribution in the model, these authors proposed to replace the number of daughters per sire by daughter equivalents (DE), weights actually derived from reliabilities ($\mathbf{R^2}$) and a variance ratio (k) to traduce the precision of daughter information:

$$DE_i = \frac{kR_i^2}{1-R_i^2} \text{ [5]}$$

Animal model drastically increases the number of equations compared with sire models to increase computational difficulties and time. However, animal model has interesting properties, as described by Kennedy *et al.* (1988). These authors showed that, when used with the complete additive genetic relationship matrix, the animal model can account for changes in genetic mean and variance due to drift, non random mating or selection.

- **From single trait to multi-trait analyses**

Selection is actually based on a combination of several traits of interest. It is more realistic to simultaneously evaluate animals on several traits of interest. Multi-trait analyses use the phenotypic and genetic correlations between these traits. The first use of BLUP for multi-trait evaluation was described in 1976 (Henderson and Quaas).

The model for the analysis of two traits (*1* and *2*) can be written as:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix}\begin{bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\end{bmatrix} + \begin{bmatrix}\mathbf{Z}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_2\end{bmatrix}\begin{bmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{bmatrix} + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix} \text{ [6]}$$

G is the matrix of additive genetic variance-covariance and is written as:

$$\mathbf{G} = \text{var}\begin{bmatrix}\mathbf{a}_1\\ \mathbf{a}_2\end{bmatrix} = \begin{bmatrix}\mathbf{g}_{11}\mathbf{A} & \mathbf{g}_{12}\mathbf{A}\\ \mathbf{g}_{21}\mathbf{A} & \mathbf{g}_{22}\mathbf{A}\end{bmatrix} \text{ [7]}$$

with g_{11} and g_{22} being the additive genetic variance for traits *1* and *2* and $g_{12} = g_{12}$ the additive genetic covariance between the two traits. **R** is the matrix of residual variance-covariance and is written as:

$$\mathbf{R} = \text{var}\begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix} = \begin{bmatrix}\mathbf{r}_{11}\mathbf{I} & \mathbf{r}_{12}\mathbf{I}\\ \mathbf{r}_{21}\mathbf{I} & \mathbf{r}_{22}\mathbf{I}\end{bmatrix} \text{ [8]}$$

where r_{11}, r_{22}, r_{12}, r_{21} are the variances and co-variances for residual effects.

It follows that the HMME for multi-trait analyses are:

$$\begin{bmatrix}\mathbf{X'R^{-1}X} & \mathbf{X'R^{-1}Z}\\ \mathbf{Z'R^{-1}X} & \mathbf{Z'R^{-1}Z} + \mathbf{A^{-1}} \otimes \mathbf{G^{-1}}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{b}}\\ \hat{\mathbf{a}}\end{bmatrix} = \begin{bmatrix}\mathbf{X'R^{-1}y}\\ \mathbf{Z'R^{-1}y}\end{bmatrix} \text{ [9]}$$

where $\otimes$ is the Kronecker product between 2 matrices (Searle, 1982), and:

$$\mathbf{X}=\begin{bmatrix}\mathbf{X}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{X}_2\end{bmatrix},\ \mathbf{Z}=\begin{bmatrix}\mathbf{Z}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_2\end{bmatrix},\ \hat{\mathbf{b}}=\begin{bmatrix}\widehat{\mathbf{b}_1}\\ \widehat{\mathbf{b}_2}\end{bmatrix},\hat{\mathbf{a}}=\begin{bmatrix}\widehat{\mathbf{a}_1}\\ \widehat{\mathbf{a}_2}\end{bmatrix}\ \text{and}\ \mathbf{y}=\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix}\ [10]$$

Again, multi-trait analyses increases accuracy of breeding value estimation. The gain from using correlated traits is partitioned into a direct gain from measuring other traits and a gain because fixed effects are estimated more precisely. Use of residual covariance between traits allows for better data connectedness (Thompson and Meyer, 1986).

- **From national to international evaluations**

Throughout the 1980s, export of North American Holstein semen to other countries became widespread (Powell and Norman, 2006). These bulls had a large number of daughters in several countries and were evaluated in several places. For a sire, at an early stage of its productive life, combining pedigree and performances at the international level is known to increase the accuracy of its breeding value estimates in the evaluation system of the importing country. Moreover, the need for a tool for international comparisons was an important issue for fair exchange of semen and for genetic improvement programs.

In 1994, Interbull started to combine national genetic evaluations across countries and provided evaluations for all bulls on each participating country's scale with 4 Nordic countries and 2 breeds.

In August 1995, the Multiple Across Country Evaluation (**MACE**) proposed by L. R. Schaeffer (1994), was adopted. This sire model for multi-trait analysis considers each trait evaluated in a country as a different trait from the one evaluated in the other countries. Therefore, different levels of heritability among countries and genetic correlations between countries less than one are used. In fact, the genetic correlations account for differences among countries in national models for data collection and genetic evaluations and for genotype by environment interaction. Finally, this model assumes a zero residual co-variances between countries as daughters are supposed to be recorded in only one country.

The MACE model requires:

- an international relationship matrix including sires and maternal grand-sires born in all participating countries. This pedigree is currently evolving toward a sire-dam pedigree (Jakobsen and Fikse, 2009).
- de-regressed proofs (**DRP**) of bulls as observations (Sigurdsson and Banos, 1995). DRP are a vector of pseudo-performances derived from national estimated breeding values and free from pedigree information. DRP are used as proxies of DYD.
- Effective daughter contributions (**EDC**) (Fikse and Banos, 2001) to weigh these phenotypes instead of the number of daughters of each sire. EDC consider together the contemporary group structure, the correlation between repeated records, and the reliability of dams:

$$EDC_i = \sum_j \frac{\mathrm{kR_j}(o)}{\text{4-R}_{\mathrm{j(o)}} \cdot (1 + R_{dam}(o))} \quad [11]$$

Summation is over all daughters j of a sire i, $k = \frac{4 - h^2}{h^2}$ and $R_{j(o)}$ is the reliability of animal j's own record, $R_{dam(o)}$ is the reliability of the dam of j's own performance.

The MACE model for DRP submitted by country *c* (i.e., provided to Interbull) is usually denoted:

$$\mathbf{y_c} = \mathbf{1}\mu_c + \mathbf{Z_i Q w_c} + \mathbf{Z_c s_c} + \mathbf{e_c} \quad [12]$$

where μ_c is the mean effect for country *c*, $\boldsymbol{w_c}$ is a vector of genetic effect for groups of unknown parents (Quaas, 1988), **Q** is the matrix that relates sires to these groups. The latter are defined for unknown sires and maternal grand-sires on the basis of their country of origin, year of birth of their progeny and selection paths. Given two countries (*1* and *2*) and *n* bulls, the genetic variance-covariance matrix is:

$$\mathbf{G} = \mathrm{var}\begin{bmatrix}\mathbf{s_1}\\ \mathbf{s_2}\end{bmatrix} = \begin{bmatrix} g_{11}\mathbf{A}_{nn} & g_{12}\mathbf{A}_{nn} \\ g_{21}\mathbf{A}_{nn} & g_{22}\mathbf{A}_{nn}\end{bmatrix} [13]$$

Given $\mathbf{D_i}$ for the diagonal matrix of EDC in country *c*, the matrix of residual variance-covariance is:

$$\mathbf{R} = \mathrm{var}\begin{bmatrix}\mathbf{e_1}\\ \mathbf{e_2}\end{bmatrix} = \begin{bmatrix} \sigma^2_{e1}\mathbf{D_1} & 0 \\ 0 & \sigma^2_{e2}\mathbf{D_2}\end{bmatrix} [14]$$

Following the manipulation of HMME proposed by Quaas (1988) to more easily account for genetic groups of unknown parents, the HMME for multi-trait analysis are (new parts appear in red):

$$\begin{bmatrix} \mathbf{X'R^{-1}X} & \mathbf{X'R^{-1}Z} & \mathbf{0} \\ \mathbf{Z'R^{-1}X} & \mathbf{Z'R^{-1}Z + A^{-1} \otimes G^{-1}} & \mathbf{-A^{-1}Q \otimes G^{-1}} \\ \mathbf{0} & \mathbf{-Q'A^{-1} \otimes G^{-1}} & \mathbf{Q'A^{-1} \otimes G^{-1}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{c}} \\ \mathbf{Q}\hat{\mathbf{w}} + \hat{\mathbf{a}} \\ \hat{\mathbf{w}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'R^{-1}y} \\ \mathbf{Z'R^{-1}y} \\ \mathbf{0} \end{bmatrix} [15]$$

In MACE implementation, all relatives, whatever their country of origin, are considered. MACE breeding values especially benefit from good connectedness of the data due to common bulls across countries, i.e., bulls with daughters in several countries, to accurately estimate the country effects.

- **From single trait MACE to multi-trait MACE**

Selection is often based on combined estimated breeding values and especially on total merit indices which are a combination of several breeding values estimated by a multi-trait analysis and weighed by economic values defined at national level. Traits with low heritability are often estimated together with several predictors of higher heritability. EBV for complex traits

are often predicted using traits that are easier to measure. EBV for time-dependent traits can be estimated considering traits measured at different ages as different traits.

To benefit from the sophisticated but heterogeneous national models and the international level of information, the interest in multi-trait analyses for multiple across country evaluation (MT-MACE) is high. This would be particularly relevant for complex traits of high economic interest such as fertility. Such analyses are currently under development (Nilforooshan, 2011) but are not yet routinely implemented.

2-1.2. THE USE OF A NEW TYPE OF INFORMATION, THE GENOMIC ONE

Besides phenotypic observations, a new kind of data - genotypes at molecular markers - can now contribute to the estimation of breeding values: using dense SNP chips, allelic variations at marker loci along the genome are known for each genotyped individual and can be used as proxies of allelic variations in genes nearby.

- **The contribution of genomic information to selection decisions**

Selection decisions depend on the knowledge of an estimated parent average (**PA**) and an estimated Mendelian sampling contribution (**MS**) at the time of selection. The use of genome wide approaches is a clear turning point in genetic evaluations as it permits an early estimation of MS terms.

Under classical BLUP approaches, the accuracy of PA estimates is increased by the quantity of information available, i.e., by adding information on the parents themselves or on relatives or on correlated traits. Implementing BLUP under animal model, multi-trait analyses or international evaluations served this objective.

The accuracy of the estimated MS term, using BLUP, is increased by using the individual's own phenotypic record or progeny information, which are obtained relatively late in the life of the animal, i.e., after progeny testing. In contrast, molecular markers can be used to immediately estimate MS contribution at birth (or even before). Meuwissen *et al.* (2001) demonstrated that genomic evaluation can increase the accuracy of EBV from the increased accuracy of the MS term estimate. Daetwyler *et al.* (2007) emphasized that the exploitation of MS variation is the major source of increased genetic progress of genome-wide approaches over BLUP approaches.

However, including genomic information into genetic evaluations is not as straightforward as including a larger amount of data: the BLUP model cannot be adapted as in the past. New types of evaluation models have been especially developed to fit this information.

- **Development of genomic evaluations: principle**

Until recently, DNA markers have only been used to search for quantitative trait loci (**QTL**) in the genome to use them in marker assisted selection (**MAS**). With the increased density of molecular markers, it becomes possible to find markers in linkage disequilibrium with any QTL. Moreover, because of the knowledge of genome-wide markers, all QTL can be

considered simultaneously. Two approaches have been developed to perform genomic predictions:

- Genomic breeding values are derived from prediction equations established in a reference (or training) population:

The method was first proposed by Meuwissen *et al.* (2001), further developing older ideas, e.g., (Lande and Thompson, 1990). It conceptually involves two steps (Hayes *et al.*, 2009):

- The entire genome is implicitly divided into small segments. Their effects are estimated in a reference population (**RP**) in which animals are both phenotyped and genotyped. Prediction equations based on these estimations are then derived from the effects of all segments.
- Animals of interest such as selection candidates can be genotyped and their genomic breeding values are predicted as the sum of the effects of the chromosome segments they carry.

Linear and non linear methods can be used to estimate the chromosome segment effects. Among others, least square regression, random regression BLUP, on the one hand, or Bayesian models, on the other hand, are often cited in the literature (de Roos, 2011) . The main difference between these models lies in the assumed distribution of marker effects.

- Genomic breeding values are derived from a G- BLUP approach

This alternative strategy consists in computing estimated breeding values from the usual mixed model equations where the relationship matrix between genotyped animals contains relationship coefficients estimated from the similarity between the observed genotypes. BLUP using such a genomic relationship matrix is called **G-BLUP** (Van Raden, 2008) and can be shown to be equivalent to BLUP on marker effects.

Practical applications of genomic evaluations have shown that linear methods based on BLUP are relatively easy to implement in terms of programming and computing requirements. But Bayesian methods tend to outperform them for traits influenced by a few QTL with large effects (Daetwyler *et al.*, 2010).

Genomic predictions raise a major statistical problem related to the relatively large number of effects to estimate compared with the number of observations. Other methods were therefore developed to reduce the number of SNP by identifying the most informative markers: the Elastic-Net (Croiseau *et al.*, 2011), the Bayesian Lasso (Legarra *et al.*, 2011), the sparse PLS (Colombani *et al.*, 2011) methods among many others were studied for this purpose.

- **Dependency between genomic and classical evaluations**

Genomic evaluations will not replace the evaluations based exclusively on phenotypes. The quality of genomic evaluations closely depends on phenotypes which are analyzed through classical evaluations based on sophisticated genetic models.

In fact, the phenotypic and pedigree-based evaluations provide observations which are used for the estimation of SNP effects in the reference populations. Corrected performances such

as daughter yield deviations or de-regressed proofs can be used as input data for genomic predictions. Moreover, multi-national RP were shown to be beneficial to increase the reliability of genomic evaluations (Lund *et al.*, 2010, Lund *et al.*, 2011, Van Raden *et al.*, 2009). DRP from international genetic evaluations are therefore required as phenotypes. In fact, genomic evaluations are dependent on the national but also international genetic evaluations which rely on BLUP or MACE analyses.

Moreover, in animals (Simianer *et al.*, 2011b) as well as in humans (Maher, 2008), genomic information was shown to only explain a part of the additive genetic variance, the unexplained part being referred to as the "missing heritability" part. Since genetic markers do not explain 100% of the genetic variance, genomic information has to be combined with polygenic information to estimate the complete genetic effects. When genomic information is combined with the polygenic one to compute estimated breeding values, EBV are called "genomically enhanced breeding values" and abbreviated by **GEBV**.

2-1.3. COMBINATION OF GENOMIC AND POLYGENIC INFORMATION: HOW TO GET GEBV?

To increase the quality of the genetic evaluation process, it is not only relevant to increase the quality of genomic predictions of genotyped animals with polygenic information (as seen previously) but also to increase the quality of genetic predictions of non genotyped animals with genomic information of the genotyped ones.

Two alternative approaches exist to combine genomic with polygenic information. On the one hand, multi-steps approaches are currently implemented to blend all types of estimated breeding values. However, not all the population benefits from genomic information of genotyped relatives. On the other hand, a single step approach is being developed to simultaneously estimate the breeding values of genotyped and non genotyped animals whether they are phenotyped or not.

Among single and multi-step approaches, four methods are presented here according to the number of steps involved (from 1 to at least 3) and the population for which GEBV are delivered: only genotyped animals, genotyped animals and their ancestors, or all animals, whether genotyped or not.

- **Case 1: GEBV computed for genotyped animals only**

Genotyped animals are evaluated by genomic evaluation on the one hand and by classical evaluation on the other hand. Genomic breeding values and classical EBV are then combined in a post-processing phase. This approach usually requires two (if genomic predictions are based on the use of a genomic relationship matrix) to three steps (if based on the estimation of marker effects) and is therefore called a multi-step approach.

Before all, classical evaluations are run for all animals with phenotypes (C1 evaluation). The C1 EBV for the genotyped animals in the reference population are converted into phenotypes, e.g., de-regressed proofs or DYD.

Based on the knowledge of genotypes and phenotypes, genomic effects (i.e., marker or breeding value) can be estimated. Direct genomic value or **DGV** for the genotyped animals are actually computed either from the prediction equations or from a G-BLUP approach (G2 evaluation).

G2 evaluations include less pedigree information than the C1 ones: there are fewer (genotyped) ancestors included in G2. To account for this difference, a third evaluation is run, i.e., a classical evaluation including only genotyped animals and excluding all non genotyped animals such as ancestors and dams (C3 evaluation). Parent average (PA) of genotyped animals can be estimated from C1 and C3.

Finally, the genomic information from DGV is blended with the pedigree information from PA into one final GEBV using selection index theory. This strategy was described by Goddard and Hayes (2007) and implemented by Van Raden *et al.* (2009) for North American Holstein bulls. The three types of EBV for genotyped animals are weighed by coefficients derived from selection index theory which are function of heritability and accuracies of C1, G2 and C3.

Such an approach benefits to genotyped animals and enhances the quality of their genetic evaluations by combining polygenic information with extra genomic information. However, non genotyped animals do not benefit from the genomic information of their genotyped relatives.

This approach requires 3 evaluation runs and a post-processing treatment to blend all sources of information. Despite the number of steps, this indirect approach was said to be easy to implement using existing software. Another advantage is that only one final value is delivered and expressed on the same scale as classical EBV which is therefore easy to use and to understand by breeders. Such approach is the most frequently implemented one but it is still an approximation and may lead to biased GEBV and overestimated reliabilities (Van Raden *et al.*, 2009)

- **Case 2: GEBV computed for genotyped animals and their ancestors**

Genotyped animals and their non genotyped ancestors can be evaluated in a same evaluation. In fact, genotypes of ancestors should be inferred. The additive genetic value is computed as the sum of genomic and polygenic effects for both groups of animals.

Genomic information from actual molecular marker information (for the genotyped animals) and from inferred molecular marker (for the non genotyped ancestors) are analyzed together with the polygenic information of the ancestors. This strategy has been implemented in France based on Marker-Assisted Selection (MAS) combined to genome-wide information. It is called the G-MAS approach and was described by Boichard *et al.* (2010) at the 9th WCGALP in Leipzig.

The G-MAS approach relies on a preliminary step which aims at identifying the informative haplotypes using two ways:

- QTL detection based on Linkage Disequilibrium and Linkage Analysis (LDLA) using haplotypes of consecutives SNP (Druet *et al.*, 2008, Meuwissen and Goddard, 2001);
- SNP pre-selection by Elastic-Net (**EN**) (Croiseau *et al.*, 2011).

It is known that markers and QTL only explain a fraction of the additive genetic variance and the respective weights for polygenic and haplotypic effects have to be given depending on variance components (Guillaume *et al.*, 2008): the variance explained by QTL and SNP is assumed to be between 40 and 60% depending on traits. The variance of each QTL haplotype is estimated by LDLA. All SNP haplotypes detected by EN are assumed to have the same variance.

Before genomic evaluations are performed, missing genotypes and the different marker haplotypes at each QTL need to be inferred whenever it is possible for each evaluated animal.

The first step of the genomic prediction process is the computation of phenotypes (DYD or DRP) from the most recent and most complete classical evaluation (previously called C1).

In a second step, haplotypic and polygenic effects are jointly estimated in a BLUP-based evaluation (G2). The total breeding value of each animal is then computed as the sum of all QTL effects and the "residual" polygenic component. Using the notations of Guillaume *et al.* (2008), the model is written as followed:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \sum_{i=1}^{n_{QTL}} \mathbf{Z}_{v_i} \mathbf{v}_i + \mathbf{e} \quad [16]$$

where $\mathbf{y}$ is the vector of phenotypes; $\boldsymbol{\beta}$ is the vector of fixed effects, $\mathbf{u}$ is the vector of random polygenic effects, $\mathbf{v_i}$ is the vector of random haplotypic effect, e.g., for QTL, $\mathbf{e}$ is the vector of random residual errors ; $\mathbf{X}$, $\mathbf{Z}$ and $\mathbf{Z}_{v_i}$ are incidence matrices which associate animals to fixed, polygenic and haplotypic effects.

In this case, genotyped animals naturally benefit from the polygenic information and non genotyped ancestors benefit from the genomic information of all relatives included in the relationship matrix in G2.

- **Case 3: GEBV for all animals, genotyped or not**

Two strategies were proposed to analyze simultaneously and more broadly all genotyped and non genotyped animals:

- A bivariate approach:

Gianola *et al.* (2006) proposed to use a bivariate model to simultaneously analyze traits of genotyped and non genotyped animals. They assumed that the breeding value of a genotyped animal is the sum of polygenic and genomic effects and the genomic information is considered as a trait correlated with the polygenic information. Effects are estimated by HMME. However, genomic effects are not estimated for non genotyped animals, only the polygenic effect is estimated. The two-trait linear model was written as:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix}=\begin{bmatrix}\mathbf{W}_1\\ \mathbf{W}_2\end{bmatrix}\boldsymbol{\beta}+\begin{bmatrix}\mathbf{Z}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{Z}_2\end{bmatrix}\begin{bmatrix}\mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix}+\begin{bmatrix}\mathbf{0}\\ \mathbf{T}(h)\end{bmatrix}\boldsymbol{\alpha}+\begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix}\quad[17]$$

y is the vector of phenotypic data (e.g., any trait measurements): $\mathbf{y_1}$ consists of records of non genotyped animals, $\mathbf{y_2}$ includes records of genotyped animals. **β** is the vector of fixed effects, **u** the vector of additive genetic effects independent from the marker effects **α.** In fact, **u** is a strictly additive polygenic effect. For a given smoothing parameter *h*, **T(*h*)** is estimated in a previous step by semi-parametric procedures (not detailed here). It is assumed that:

- $\begin{bmatrix}\mathbf{u_1}\\ \mathbf{u_2}\end{bmatrix}\sim\mathbf{N}\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix},\begin{bmatrix}\mathbf{A}\sigma^2_{u_1} & \mathbf{A}\sigma_{u_{12}}\\ \mathbf{A}\sigma_{u_{12}} & \mathbf{A}\sigma^2_{u_2}\end{bmatrix}\right)$ [18]
- $\boldsymbol{\alpha}_j\sim N(\mathbf{0},\sigma^2_\alpha)$;
- $\mathbf{e}_1\sim N(\mathbf{0},\mathbf{I}\sigma^2_{e_1})$ and $\mathbf{e_1}$ is independent from $\mathbf{e_2}$: $\mathbf{e}_2\sim N(\mathbf{0},\mathbf{I}\sigma^2_{e_2})$.

$\sigma^2_{u_1}$ and $\sigma^2_{u_2}$ are additive genetic variances in individuals without and with molecular information, respectively and $\sigma_{u_{12}}$ is their additive genetic covariance. σ^2_α is the genetic variance of marker effects.

Once more, non genotyped animals do not fully benefit from the genomic information of their genotyped relatives: $\mathbf{u_1}$ is correlated with $\mathbf{u_2}$ but not with **v**.

➢ A single step approach:

Two research groups proposed an alternative method to simultaneously evaluate all genotyped and non genotyped animals in a single step approach. On the one hand, Christensen and Lund (2010) and on the other hand, Misztal *et al.* (2009), Legarra *et al.* (2009) and Aguilar *et al.* (2010) also suggested to predict the additive genetic value of all animals as the sum of polygenic and genomic effects. In contrast with Gianola *et al.* (2006), genomic effects for the non genotyped animals are inferred from genotyped animals using the pedigree-based relationship matrix as regression coefficients. In fact, the genomic information from genotyped animal is transmitted to non genotyped animals through the relationship matrix: if $\mathbf{a_1}$ and $\mathbf{a_2}$ are the additive genetic values of non genotyped and genotyped animals combining polygenic and genomic information and σ^2_a is the genetic variance (Legarra *et al.*, 2009):

$$p(\mathbf{a}_1\,|\,\mathbf{a}_2)\sim N(\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{a}_2,(\mathbf{A}^{11})^{-1}\sigma^2_a)\quad[19]$$

The approach especially relies on a modification of the numerator relationship matrix **A** into a matrix **H** by replacing in the pedigree based relationship matrix **A** the part corresponding to genotyped animals by the genomic relationship matrix **G** (Misztal *et al.*, 2009):

$$\mathbf{H}=\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{G}\end{bmatrix}=\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}+(\mathbf{G}-\mathbf{A}_{22})\end{bmatrix}\quad[20]$$

where subscripts *1* and *2* represent ungenotyped and genotyped animals and **G** is a genomic relationship matrix. This **H** matrix replaces the usual relationship matrix in a BLUP evaluation of all genotyped and non genotyped animals. But in HMME, it is the inverse of **H** which is needed. Both research groups discovered that in fact, $\mathbf{H}^{-1}$ has a simple form, which was rather unexpected:

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \quad [21]$$

To avoid potential problem due to singularity with the inversion of **H**, **G** can be modified as:

$$\mathbf{G} = w\mathbf{G}_r + (1-w)\mathbf{A}_{22} \quad [22]$$

where *w* can be interpreted as the relative weight of the polygenic effect needed to explain the total additive variance (Christensen and Lund, 2010), $\mathbf{G}_r$ is the original matrix constructed from genotypes before including weights.

Single step procedures are therefore considered as a natural way to combine all types of information on all animals. However, there are difficulties to implement single step procedures for large numbers of genotyped animals, to adapt to complex genetic models and to consider foreign genotyped bulls without domestic daughters.

Conclusion: From sire to animal model, from single to multi-trait analyses, from national to international evaluations, the genetic predictions have been enhanced by including an increasing amount of information. Information on relatives, on correlated traits at the national and then at the international level has been continuously added. With genomic information, a new kind of evaluation has emerged which now needs to be combined with the classical sources of information. Multi-step approaches are already implemented but the entire population do not benefit yet from the availability of genomic information. Single step approaches are promising but require heavy software adaptations.

2.2. New breeding strategies to use genomic information

2-2.1. PRINCIPLE AND LOGISTICS OF A REGULAR PROGENY TEST SCHEME

Progeny testing has been intensively used for decades in dairy cattle breeding schemes to identify bulls of high genetic merit. Given breeding objectives and parent information, a sample of bulls is randomly mated to cows to get daughters with performances in different environments. After progeny testing, based on more accurate PA and MS estimations, full-sib or half-sib bulls can be ranked.

Among others, Schaeffer (2006) precisely described the logistics for progeny testing. Elite females are identified each year to become dams of young bulls and these are mated to the best sires from the current generation, the sires of bulls. These sires are highly selected and can be imported from foreign breeding schemes. Young bull calves are purchased by breeding companies and moved to studs. Once young bulls are sexually mature, i.e., at 1 year of age or a bit more, they are mated to a sample of cows in the population in order to produce a minimum number of daughters (e.g. 100) recorded for different trait groups such as production, conformation, functional, calving ease, etc… Once these daughters have completed their first lactations, the young bulls get their first EBV. For production traits, the reliability of EBV is usually about 75% or more. At this point, the young bull is "proven": it may be culled or returned to service. The entire process of progeny testing takes about 5 to 6 years.

2-2.2. MOVING AWAY FROM PROGENY TESTING SCHEMES

- **Several designs to implement genomic selection**

Three main strategies can be envisioned to integrate a genomic selection step in dairy breeding schemes.

- Strategy I: genomic selection as a pre-selection tool

The first strategy consists in selecting young bull calves based on their genomic breeding values before entering regular progeny testing.

Marker assisted selection was already envisioned as such a pre-selection tool as it provided a smaller gain in accuracy than genomic selection. Such a strategy was implemented in France between 2001 and 2008 with the first Marker-Assisted Selection (MAS) program based on microsatellites markers and between 2008 and 2009 with the second MAS program based on SNP markers (Fritz and Guillaume, 2008).

Currently, some AI companies have decided to reinforce their progeny testing program together with the implementation of a genomic pre-selection step (e.g., ABS breeding company in the USA). This is a way to increase selection intensity while at least keeping the

same level of reliability as progeny testing provided for the choice of elite sires. This is also a way to get both GEBV and EBV for all sires.

- Strategy II: disuse of progeny testing schemes

In the second strategy, genomic selection may completely replace progeny testing to identify bulls with high genetic merit and to take final selection decisions. In some countries, this conversion is already well advanced (France, Germany) and some bull barns were closed or will be used to house more young bulls (Ducrocq and Santus, 2011).

However, two options can be recognized: the selected young bulls are used for service either with (1) or without (2) restriction. According to the first option, young bull semen are distributed for a short period of time or for a restricted number of mating. Some AI companies, e.g., Semex in Canada, CRV in the Netherlands or CREAVIA in France, among others, have decided to limit AI services. In this case, sires are still being progeny tested but through commercial sales rather than through an established progeny testing scheme (Ducrocq and Santus, 2011). It mainly aims at buffering the risks due to lower GEBV reliabilities and it favors random mating for better data recording and breeding value estimation.

- Strategy III: disuse of national breeding programs

Finally, some countries may abandon their own breeding program in the future and simply import bulls based on prediction equations derived on their local conditions (Goddard and Hayes, 2007).

2-2.3. OPTIMIZATION OF BREEDING SCHEMES BASED ON GENOMIC INFORMATION

Genomic tools are now used to optimize breeding schemes. Various key-elements might be depicted to benefit from the advantages of genomic selection over progeny testing:

- **Management of genetic gain**
  - Reduction of generation interval

Because of information on MS term, GEBV at birth were shown to be far more accurate than classical EBV at birth, i.e., parent average (PA). Van Raden *et al.* (2009) reported an average gain in reliability of 23% for genomic predictions compared with published PA over 26 traits from low to moderate heritability. Gains are consistent with those observed in France in Holstein breed based on either national or multi-national reference population: averaged over four traits of interest, reliability of GEBV was respectively 20% and 29% higher than the reliability of PA (Lund *et al.*, 2011).

It follows that GEBV accuracies of young bulls are almost as high as accuracies after progeny testing. In France, bulls used to be proven with an objective of a minimum reliability of 50% for functional traits and 70% for production traits even though the final values were

usually a bit higher. With genomic evaluations, a 60% reliability for fertility and a 70% reliability for milk yield can be reached among young bulls (Fritz *et al.*, 2010).

Selecting young bulls based on GEBV may dramatically reduce the generation interval on the sire of cow pathway (from 5.5 to 2 years) with only a small loss in accuracy. The generation interval can not only be reduced for young bulls but also for sires of bulls (Schaeffer, 2006). On the female side, the possibilities of reduction are small because, at least in France, most cows selected as bull dams are heifers.

- Increase of selection intensity

Further genetic gains can be made by increasing selection intensity, not only on the male but also on the female pathways: elite bull dams and large numbers of calves of both sexes can be genotyped to select few of them. It is even possible to genotype and genomically evaluate embryo and select the best ones among them before transfer (Le Bourhis and Humblot, 2010).

- Reduction of breeding costs

Schaeffer (2006) compared a traditional progeny testing program with a genomic selection scheme under a typical Canadian-like dairy cattle situation. The logistic costs of the breeding scheme were reduced by 92 % when progeny testing was abandoned in favor of genomic selection schemes. At the same time, genetic gain per year was doubled.

In fact, genomic selection offers the potential to dramatically increase genetic gain and decrease breeding costs (especially in the case of strategy II). Cost savings could then be invested to even intensify selection by genotyping a larger number of animals on the different pathways.

- More balanced objectives of selection

In the past, large gains were made on production traits. Genomic selection would now help to select more efficiently on traits with lower heritability, e.g., among functional traits. Selection objectives could be more diversified but only if dedicated weights on, e.g., functional traits, in Total Merit Index are increased. For example, this trend is currently observed throughout Europe, TMI in France will be reconsidered from February 2012.

- **Management of inbreeding**

Molecular markers offer new tools to manage inbreeding. Bouquet *et al.* (2011) showed that the use of dense marker information improves the estimation of inbreeding coefficients compared with those based on pedigree information. Monitoring inbreeding rate will be especially required. Many studies have shown that genomic strategies may affect inbreeding rate in different directions:

- if generation intervals stay the same as in progeny testing scheme (strategy I), genomic selection could result in lower rates of inbreeding (Daetwyler *et al.*, 2007). By promoting across family selection, the selected young bulls could be chosen among more families and would be less related (Simianer *et al.*, 2011a).

- if generation intervals decrease (strategy II), the level of inbreeding could be lower per generation but higher per year. In fact, inbreeding rate can be multiplied by up to three (de Roos, 2011, Simianer *et al.*, 2011a).

Colleau (2009) simulated several schemes with data from the Montbéliarde breed to optimize genetic gain relative to inbreeding. Compared to a classical progeny testing scheme, the increase of genetic trend ranged from 72 to 100% whereas the increase of inbreeding ranged from -23% to +100% depending on the scenario.

Implementing genomic selection could lead to a higher genetic gain and a lower inbreeding rate. According to Colleau (2009), limited or no use of highly selected proven sires as well as a sharp increase of the number of sires of sons would decrease the inbreeding rate.

2-2.4. CHANGES IN SEMEN DISTRIBUTION

Compared with the EBV accuracy of bulls with progeny, the accuracy of GEBV is lower and the confidence interval of GEBV is increased. GEBV are more likely to change with additional information than EBV obtained after progeny testing. Consequently, rankings based on GEBV are less precises than rankings based on classical EBV as commonly used before.

To account for breeding value uncertainties at breeding time and to promote a better balance in selection objectives, the AI industry, e.g. in France, tends to distribute bull semen as packs rather than as individual bull. The packs contain semen straws of several young bulls. They are grouped according to selection objective or price or genetic diversity among other criteria. This gives less importance to individual breeding values and therefore buffers risk due to lower reliabilities.

More bulls are now used for AI but for a shorter period of time. The semen availability per bull is also decreased mainly because of its young age at time of semen production and dissemination.

Finally, the earlier availability of GEBV but their lower accuracy may lead to shorter bull productive life and may reduce importance given to top sires in the international lists.

<u>Conclusion:</u> Genomic information is now available. The statistical methods for genetic predictions have been modified to deliver genomically enhanced breeding values. GEBV have interesting properties which suggest modifying the organization of breeding programs to optimize genetic gains. Using genomic breeding values as a pre-selection tool before progeny testing or, more radically, removing this long process from breeding programs, is a huge change which has already impacted the semen distribution. Great changes are also expected in the global process of genetic evaluation: approaches are currently developed to combine classical and genomic evaluations.

CHAPTER 3 - The reasons for bias in classical evaluations after a genomic selection step

The role of progeny testing in dairy cattle breeding programs is no longer the same with the availability of early and accurate predictions such as GEBV. It may not only affect breeding designs and strategies for semen distribution but also the quality of classical evaluations. In this chapter, cases and conditions for bias in breeding value estimation are described to better understand why genomic selection may affect national evaluations.

3.1. Best and unbiased genetic predictions

3-1.1. DEFINITIONS OF BIAS AND PRECISION OF AN ESTIMATOR

A desirable property of a parameter estimator is to be unbiased. Otherwise the true value of the parameter is **systematically under- or overestimated** by the estimator. In other words, if there are n repeated samplings from a given distribution, the estimator should have, on average, the correct value.

Assessment of the quality of an estimator T_n is based on the difference between T_n and the true value, θ. There are two sources of error: a "random" one, i.e., $T_n - E(T_n)$, and a systematic one, i.e., $E(T_n) - \theta$, which is called bias. Finally, the estimation error can be broken down into:

$T_n - \theta = [T_n - E(T_n)] + [E(T_n) - \theta]$. [23]

An estimator is unbiased when $E(T_n) - \theta = 0$, i.e., $E(T_n) = \theta$. The quality of an estimator depends on the quantity and quality of information included in the estimation process, its preision can be assessed by:

- its accuracy which is usually computed as the correlation between the true and estimated values , i.e., $\rho(\theta, T_n)$.
- the predictor error variance (**PEV**) which measures the dispersion of the estimates around the true value. This is the variance of the difference between true and estimated values: $PEV=var(\theta - T_n)$.

Finally, the mean squared error (**MSE**) captures both the error and precision of an estimator: $MSE = E(T_n - \theta)^2 = Var(T_n) + [E(T_n) - \theta]^2$. [24]

Applied to livestock evaluations where the true value to estimate (or "predict" if we use Henderson's terminology) is the breeding value (a), measures of error and precision of breeding value estimates ($\hat{a}$) are detailed in Table 1.

Table 1: Measures of error and precision in the estimation of breeding values

Statistical indicators		Definition
Measures of error	Random part of the estimator error	$\varepsilon = \hat{a} - E(\hat{a})$
	Systematic part of the estimator error , i.e., the bias	$\delta = E(\hat{a}) - a$
Measures of precision	Accuracy	$R = \rho(a, \hat{a})$
	Reliability or coefficient of determination	$R^2 = \rho^2(a, \hat{a})$
	Prediction error variance	$PEV = \text{var}(\hat{a} - a)$
Combined measure of error (i.e., bias) and precision	Mean squared error	$MSE = Var(\hat{a} - a) + [E(\hat{a}) - a]^2$

3-1.2. STATISTICAL PROPERTIES OF BLUP SOLUTIONS

Applied to animal breeding, the BLUP method provides predictions of random effects, e.g., animal breeding values â and of residuals ê, and estimates of fixed effects such as environmental effects $\hat{\beta}$. They are linear functions of observations with desirable statistical properties:

- BLUP provides unbiased predictions of animal breeding values: $E(\hat{a}) = E(a)$;
- BLUP provides the best predictions of breeding values, under the constraint of unbiasedness. The method maximizes the correlation between true (a) and predicted breeding value (â) and minimizes PEV.

However, the true reliability can not be computed but reliability is usually derived from the corresponding diagonal elements of the inverse of the HMME coefficient matrix $\mathbf{A}^{-1}$. Due to computational constraints, this reliability is often further approximated for example using the information source method developed by Harris and Johnson (1998). The latter successively combines reliabilities of independent information provided by parents, the animal themselves and progeny.

3-1.3. ASSUMPTIONS FOR OPTIMAL PROPERTIES OF BLUP

Linear models assume that **sampling of the data is random** and that the **model describing the data is correct,** i.e., it includes all significant factors of variations. Additional conditions are usually assumed to benefit from the optimal BLUP properties (Henderson, 1975).

First, **the infinitesimal model** where the additive genetic value is assumed to be the sum of a very large (infinite) number of loci each with very small effects, is supposed.

The distribution of the random (a and e) and observed (y) variables should be multivariate normal and selection should be based on a linear function of the records.

The genetic variance in the base population (animals without known parents), σ_a^2, is assumed to be known. Base animals are supposed to be unselected (breeding values of the base population have an expected value of 0), unrelated and non inbred.

A very important condition is that all information on which selection is based should be included in the analysis so that **the relationship matrix A is complete**. In fact, all animals involved in the selection decisions back to the base population (Sorensen and Kennedy, 1983) must be included to compute **A**. With the animal model, through the relationship matrix, any additive genetic value can be described as a linear function of breeding values of base animals and subsequent MS contributions that are unaffected by selection (Kennedy *et al.*, 1988).

Using matrix notations, this can be summed up as: $\mathbf{a} = \mathbf{T}\varphi$, where **T** is a matrix of appropriate coefficients that relate animal breeding values to ancestral breeding values and MS terms, φ, from subsequent generations.

Bulmer (1971) showed, assuming normality, that MS is independent from the breeding values of the parents and, therefore, the distribution of MS before selection in the generation of the animal is unaffected by any form of selection occurring in previous generations. It is therefore assumed that **MS terms are normally distributed with an expected value of 0 and variance of** $\frac{1}{2}(1-\overline{F})\sigma_a^2$ where $\overline{F}$ is the average inbreeding coefficient of the animal's parents.

Finally, $E(\mathbf{a}) = \mathbf{T}E(\varphi) = 0$ and $\text{var}(\mathbf{a}) = \mathbf{A}\sigma_a^2 = \mathbf{T}\,\text{var}(\varphi)\mathbf{T}'\sigma_a^2$ with var(φ) and σ_a^2 being independent from selection. In such a case, changes in genetic variance due to selection, inbreeding or genetic drift are accounted for in the HMME (Sorensen and Kennedy, 1983).

In summary, the optimal BLUP properties strongly depend **on the completeness and correctness of information about the selection process.**

3.2. Cases of biases in genetic evaluations

3-2.1. BIASES DUE TO SELECTION IN CLASSICAL EVALUATIONS

Real conditions rarely provide complete and perfect data for breeding value estimation. The first and main reason is **selection**. It can be considered as a **non random missing data process.** Different forms of selection can be distinguished (Schaeffer *et al.*, 1998):

- ***parental selection*** (and/or assortative mating) which violates the assumption of random mating;
- ***sequential selection*** when decisions for selection are based on a correlated trait which is not included in the evaluation model: it was especially studied for weaning weight on the subsequent evaluations of animals for yearling weight in beef cattle (Pollak and Quaas, 1981, Pollak *et al.*, 1984);
- ***selective phenotyping*** refers to (sample of) phenotypes which are not representative of the entire population. This occurs when performances are recorded from selected animals only, especially when animals are culled before extra performances are recorded: examples can be found in poultry (broilers) when selected on growth before reproduction or in pigs when selection occurred within herd before performances recording in control station;
- ***selective reporting*** when performance are not reported for all animals supposed to be tested: Mallinckrodt *et al.* (1995) described the case in beef cattle when only a part of the animals with weaning data had yearling weights and birth weights reported;
- ***use of highly selected animals***, i.e., with a different average genetic level and coming from sub-populations where the history of selection is not known. This is especially the case when foreign sires are used as elite sires in breeding schemes in dairy cattle (Bonaiti and Boichard, 1995, Pedersen *et al.*, 1999).

Some types of selection such as selective phenotyping or selective reporting may especially change the null expectation of the Mendelian Sampling terms which is a basic assumption for optimal properties of animal models.

3-2.2. OTHER SOURCES OF BIAS IN CLASSICAL EVALUATIONS

- **Altered pedigree information**

Pedigree may also be altered by selective reporting or wrong reporting of information (misidentified parents or willfully falsified data) so that the relationship matrix is no longer correct.

- **Heterogeneity of variance among subclasses**

Any difference between the actual and assumed genetic and residual parameters in the model also causes bias. In fact, heterogeneity of variance among subclasses, e.g., within herds or countries, was especially observed at the international level on MACE evaluations, due to differences in selection schemes (number of daughters per bull, number of bulls sampled per year, etc…).

Heterogeneity of residual variance was shown to affect the quality of genetic evaluations: individuals coming from more variable herds but with the same expected genetic value than cows in other herds, tend to be more selected, leading to a systematic over-evaluation (Vinson, 1987).

- **Preferential treatment**

A very important but nearly uncontrollable source of bias is preferential treatment (**PT**). It can be described as any management practice which increases, e.g., production and which is applied non randomly to one or several cows within some contemporary groups (Kuhn *et al.*, 1994).

As an illustration, elite animals such as daughters of highly selected imported bulls with expensive semen or cows under multiple ovulation and embryo transfer may be prone to PT. A technique facilitating PT is the use of growth hormone to increase milk yield without official recording and without possibility to include it as a factor into genetic evaluation models (Colleau, 1989).

Preferentially treated cows are usually systematically over-evaluated because part of the environmental effects is not identified nor removed.

- **Wrong statistical model**

Any failure of the model, i.e., missing fixed effects may induce bias. Including unnecessary fixed effects does not create bias but increases PEV (Van Vleck, 1987).

3-2.3. MECHANISMS OF BIAS IN GENETIC EVALUATIONS

In each case, **the sample observations are no longer representative of the entire population**, which generates unfair comparisons. Information to understand the entire selection process is missing. Mean and variances of the evaluated animals deviate from the mean and variance of the population.

After selection, the variance of the selected sample tends to be reduced and the mean tends to be higher than in the entire population which leads to an underestimation of the best animals and an over-estimation of the worst ones. Animals with intermediate breeding values are the least affected. Sorensen and Kennedy (1984) and Weigel and Banos (1997) actually noticed that such inconsistencies between genetic parameter estimates and their true values lead to bias in the estimation of breeding values, e.g, an underestimation of the genetic variance

causes an overestimation of the breeding values of elite bulls and conversely (Weigel and Banos, 1997).

3-2.4. THE USE OF MOLECULAR INFORMATION: NEW TYPES OF SELECTION BIAS

- **Bias in the estimation of genomic effects due to selective phenotyping**

Mackinnon and Georges (1992) studied the effect of selection bias on linkage analysis for quantitative traits. Their study has shown that selection within a sample severely biases the magnitude of the Quantitative Traits Loci (QTL) effects and reduce the power with which they were detected. Similarly, selective phenotyping may therefore affect genomic predictions based on the estimation of SNP effects.

- **Bias in genomic evaluations due to selective genotyping**

Selective genotyping refers to genotyping of a selected subset of individuals. In dairy cattle, most genotyped animals or their parents have undergone strong selection. Animals having molecular information are not a random sample from the population. Ignoring this issue in the process of genomic evaluation may lead to biased predictions.

The problem of incomplete genotyping was especially addressed by Gianola *et al.* (2006), Ehsani *et al.* (2011, 2010), Vitezica *et al.* (2010).Obviously, the genomic relationship matrix is in this case computed for genotyped animals only. The latter are known to be selected but selection decisions cannot be traced back to the base population. Unlike the genetic relationship matrix used in BLUP under animal model, the genomic matrix is far from complete and cannot completely account for selection as shown by Sorensen *et al.* (1983). Via simulations, Ehsani *et al.* (2010) and Vitezica *et al.* (2010) studied the effects of selective genotyping of the reference population on genomic predictions.

Ehsani *et al.* (2010) assumed that only the progeny of selected individuals are genotyped and included in the RP at each generation. They measured the reliabilities of genomic predictions given different scenarios. Animals were either randomly genotyped or the genotyped sample was chosen among the highest, the lowest or intermediate phenotypic values or among the extreme (i.e., highest and lowest) phenotypic values. Whatever the heritability or the selection intensity, it was clear that genomic predictions were of better quality when genotyping was at random.

Vitezica *et al.* (2010) measured the bias as the difference between simulated true breeding values and GEBV: after selection of animals (based on phenotypes or on their EBV computed with a BLUP animal model) and GEBV tended to be underestimated compared to TBV.

3.3. Bias in classical evaluations after a genomic selection step

3-3.1. Genomic selection: the combined effects of "selective phenotyping" and "sequential selection"

Two reasons for selection bias are actually occurring after genomic selection.

Selective phenotyping resulting from genomic selection may also strongly affect classical evaluations. After progeny testing, daughters of all the candidates are recorded whereas after genomic selection only daughters of the selected candidates are included in HMME.

Moreover, the selection criterion which is here GEBV is neither explicitly nor implicitly included in the classical evaluation model although GEBV are obviously correlated to EBV. This has been previously defined as a problem of sequential selection.

3-3.2. Violation of the mixed model assumptions: biased BLUP solutions

The advantage of genomic evaluation over classical evaluation is to accurately and precociously estimate the Mendelian sampling (MS) contribution of an animal. Animals with a high estimated MS contribution are then preferentially retained among all candidates so that the expected MS contribution of the selected candidates clearly deviates from zero. This was also the case after the selection step based on progeny testing but selection decisions were based on BLUP solutions including the daughter performances of all selection candidates before culling.

After genomic selection, **A** is not complete. Moreover, the missing data process is not random and the null expectation of MS terms is no longer true. This violates an important assumption underlying the optimal properties of BLUP applied to animal model (Kennedy *et al.*, 1988): it is feared that estimated breeding values will be biased after genomic selection, if not accounted for in national evaluations.

3-3.3. Bias in MACE solutions

After genomic selection, only the selected candidates receive EBV based on progeny performances. The potentially biased EBV are usually sent as national proofs to Interbull to perform international evaluations based on the MACE method. Hence, this is an incomplete and possibly incorrect data set which will be sent to Interbull. Once more, the assumptions for optimal properties of mixed linear models are violated. Furthermore, data are incorrect. These both reasons may lead to biased estimated breeding values at the international level.

In other words, BLUP and MACE solutions may be altered after a genomic selection step.

3-3.4. A WORRY SHARED BY THE INTERNATIONAL COMMUNITY

The international community was warned very early (Van der Beck, 2007) about the risk of biased estimated breeding values after genomic selection. The Interbull Scientific Advisory Committee (Ducrocq *et al.*, 2008) reported that "*the magnitude of this bias needs to be assessed. If this bias is indeed important, methods to address and reduce it should be urgently derived, otherwise objective comparisons between classical breeding and genomic selection schemes will no longer be possible. This is also essential in situations where GEBV prediction equations are developed using progeny tested bull EBV as proxies to true breeding values*". Following these remarks, the bias study was splitted into three parts: 1) the assessment of bias at the national level (chapter 4); 2) the review and implementation of approaches to prevent from bias at the national level (chapter 5) and 3) the consideration of the bias propagation and its impacts at the international level (chapter 6). Methods and results will be presented.

Conclusion: Breeding values are said to be biased if they are systematically under- or overestimated. Biases were identified when information about the selection process was missing. After a step of genomic selection in dairy cattle breeding schemes, information on the culled candidates are missing in the BLUP evaluations. This process is not random and may especially affect the distribution of the Mendelian sampling terms which strongly violates the BLUP assumptions and would lead to sub-optimal BLUP properties. MACE solutions may also be affected. It seems pertinent to assess the bias magnitude, and if large, to find ways to prevent from it and avoid its propagation at the international level.

CHAPTER 4 - Bias assessment in national evaluations after a genomic selection step

Bias and accuracy need first to be measured after a genomic selection step: several possible approaches are reviewed, one strategy is proposed and used in Patry and Ducrocq (2011b).

4.1. Assessment of bias in breeding value estimates

4-1.1. MEASURES OF BIAS AND ACCURACY

- **Comparison with true values**

Bias and accuracy of EBV (Section 3-1.1 for definition) are important measures of the quality of genetic evaluations but a priori require knowing TBV. By definition, they are never known but they can be stochastically simulated. An entire breeding program can be simulated over several generations or a real data structure (phenotypes and pedigree) can be used as a starting point. Phenotypes can then be modified for the purpose of the study to mimic the issue of interest. For example, preferential treatment is often mimicked by adding an upward bias to some observations.

In such a case, bias and accuracy can directly be computed as the difference and the correlation between true and estimated breeding values.

For illustration, Hickey *et al.* (2008) estimated bias and accuracy in genetic evaluations with genetic groups by simulating sampling TBV. Colleau (1989) generated a simple breeding scheme using a Monte-Carlo method to evaluate the potential impact of growth hormone on the discrepancies between true and estimated breeding values.

Based on simulated data, any deviation of the approximate reliabilities from the true ones also reveals a loss in the quality of genetic evaluations.

- **Comparisons between evaluations under different conditions**

In the literature, two general approaches were often described to estimate bias when TBV are not available.

The first approach compares a reference (or control) situation supposed to be unbiased to a biased one. For example, Mallinckrodt *et al.* (1995) compared estimates from unaltered data to estimates from the same data sets after they were altered by selective reporting. Bias is then computed as the average difference between EBV based on the unaltered data set and EBV based on the altered data sets.

The second approach compares subsequent evaluations. Reverter *et al.* (1994) proposed three statistics to detect bias based on:

- the correlation coefficient between consecutive predictions;
- the linear regression of recent (more accurate) on previous (less accurate) genetic prediction;
- the variance of the genetic prediction difference (i.e., of recent minus previous genetic predictions).

At the international level, consistencies in rankings are especially studied by looking at the number of common sires between top lists or by computing rank correlations.

- **Importance of repeated measures of bias and accuracy**

The study of the quality of genetic evaluations often relies on simulated scenarios which are repeated to remove random error. Then, bias or accuracy can be considered as random variables and their distribution can be studied. The average bias and standard deviation of bias can be computed over replicates. A standard error, as the ratio between the standard deviation of the bias and the squared number of replicates measures the quality of the simulation process. The standard deviation of bias can be compared to the standard deviation of the breeding value to gauge the importance of the bias.

4-1.2. FACTORS INFLUENCING THE DIRECTION AND MAGNITUDE OF BIAS

Intuitively, the magnitude of the bias will depend on the degree to which the assumptions of the BLUP animal model are violated or in other words how large the deviation from basic assumptions is.

According to the literature, the magnitude of bias depends on the proportion of altered data (Mallinckrodt *et al.*, 1995) or of the proportion of culled animals without recorded performances (Pollak *et al.*, 1984, Sorensen and Kennedy, 1984). In fact, the deviation of MS terms from the null expectation will increase with the proportion of missing data.

The magnitude of bias also depends on contrasts between observed genetic standard deviation and assumed one within the model (Mallinckrodt *et al.*, 1995) or more generally, any inconsistency between the model and the actual data regarding mean or standard deviation. According to Sorensen and Kennedy (1984), a prior genetic variance smaller than the true value leads to an underestimation of the selection response and an overestimation in the reverse case.

Moreover, the importance of bias is influenced by the precision of evaluations: Mallinckrodt *et al.* (1995) observed that bias was larger when accuracy was lower. Sorensen and Kennedy (1984) also reported the role of trait heritability.

Finally, the correlation between traits (Pollak *et al.*, 1984) in multi-trait analyses (the higher the correlations, the larger the transmitted bias) and the data structure (sample size, unbalanceness and pedigree structure) were also reported as influencing bias magnitude.

These factors were identified in the literature and can now be used in simulation studies to test the sensitivity of bias after genomic selection.

4.2. Our simulation study

One particular strategy has been implemented to simulate the effects of genomic selection and to measure bias in genetic evaluations. Methods and results are presented in article I: **Patry, C. and V. Ducrocq. 2011. Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. J Dairy Sci 94:1011-1020.**

Only the outlines of the methods and the summary of the results are presented here.

4-2.1. Outlines of the implemented method

The different elements of our strategy are:

- ✓ The use of real data structure, i.e., pedigree and phenotypes observed in the Holstein breed for a national genetic evaluation;

Simulations were based on real data from France. Phenotypes and pedigree were observed for type traits which are not the main criteria for selection: foot angle (FA) and udder depth (UD). For the sake of simplicity, phenotypes were pre-corrected performances for all fixed effects except one (i.e., herd-year-season effect).

- ✓ The definition of a control situation, i.e., the progeny testing scenario;

The set of phenotypes after progeny testing was considered as complete. EBV computed after progeny testing were supposed to be unbiased. This set of data represents the CONTROL case. However, GEBV were not available and had to be simulated: phenotypes and pedigree for a fraction of the population were used to generate GEBV and TBV from a joint distribution.

- ✓ A way to mimic selection based on genomic information among young bulls: it first required to identify the cohort of young bulls, i.e., of selection candidates;

The selected candidates were distinguished from the culled candidates based on the information about MS terms, here derived from the simulated GEBV. Selection was performed within full sib families by truncation according to a defined proportion of selected candidates.

- ✓ The simulation of the effects of genomic selection;

The data structure of phenotypes and pedigree relationships was modified: once culled candidates are identified, the corresponding phenotypes were deleted. After genomic selection, genetic evaluations were only performed based on the performances of the selected candidates using a single trait BLUP based on animal model.

- ✓ The measure of bias;

Bias was measured as the difference between true and estimated breeding values. Bias was expected to be zero after progeny testing scenario. Various measures of bias after

implementation of the genomic selection step were considered. Simulations and measures were repeated 50 times to test if the error was systematic and not random, according to the definition of bias.

- ✓ The insurance of fair comparisons between the control scenario and the others;

Estimated breeding values were computed after progeny testing and after genomic selection. The number of animals in the evaluation is smaller after genomic selection, which involves another factor of variations due to the smaller amount of information involved in the estimation. In a previous step, a larger number of candidates than the number of selected candidates were generated by simulations. After genomic selection, there are as many animals as after progeny testing. If n is the number of candidates undergoing progeny testing and 1/m the proportion of selected candidates, ((n x m) –n) candidates are drawn.

- ✓ The sensitivity analysis.

Bias was measured under various conditions: two traits, i.e., two levels of heritability, two levels of information, i.e., two levels of precision for the genomic evaluations, and two proportions of missing data. It helps in validating the results and understanding bias mechanisms.

4-2.2. SUMMARY OF THE SIMULATION RESULTS

Four main conclusions can be drawn from this study of bias assessment at the national level:

- ✓ Compared to true breeding values, BLUP breeding values were strongly underestimated for the young sires kept after genomic selection. Average bias may reach one quarter to one third of genetic standard deviation. Half of this bias was transmitted to their daughters.
- ✓ Bias magnitude increased with a lower proportion of selected candidates, a lower trait heritability, and a lower precision of genomic evaluations.
- ✓ The precision of BLUP breeding values decreased after genomic selection: their reliability decreased and their MSE increased.
- ✓ Bias magnitude may be difficult to predict in the future generations when the daughters of these young sires are mated with a new generation of genomically selected young bulls.

This simulation study presents clear evidence that national genetic evaluations are biased once young sires are selected based on genomic breeding values.

Article I

Patry, C. and V. Ducrocq. 2011.

Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle.

J Dairy Sci 94:1011-1020.

J. Dairy Sci. 94:1011–1020
doi:10.3168/jds.2010-3804
© American Dairy Science Association®, 2011.

Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle

C. Patry*†[1] and V. Ducrocq*
*UMR 1313 Génétique Animale et Biologie Intégrative, Institut National de la Recherche Agronomique, 78 352 Jouy-en-Josas, France
†Union nationale des Coopératives d'Elevage et d'Insémination Animale, 149, rue de Bercy, 75 595 Paris Cédex 12, France

ABSTRACT

A genomic preselection step of young sires is now often included in dairy cattle breeding schemes. Young sires are selected based on their genomic breeding values. They have better Mendelian sampling contribution so that the assumption of random Mendelian sampling term in genetic evaluations is clearly violated. When these sires and their progeny are evaluated using BLUP, it is feared that estimated breeding values are biased. The effect of genomic selection on genetic evaluations was studied through simulations keeping the structure of the Holstein population in France. The quality of genetic evaluations was assessed by computing bias and accuracy from the difference and correlation between true and estimated breeding values, respectively, and also the mean square error of prediction. Different levels of heritability, selection intensity, and accuracy of genomic evaluation were tested. After only one generation and whatever the scenario, breeding values of preselected young sires and their daughters were significantly underestimated and their accuracy was decreased. Genomic preselection needs to be accounted for in genetic evaluation models.
Key words: selection bias, genomic selection, BLUP, dairy cattle

INTRODUCTION

In dairy cattle breeding, the recent development of genomic tools and methods has led to quick implementation of genomic selection. Due to a higher accuracy of evaluation at birth and a shorter generation interval, an increase of genetic gain is expected as well as a better management of genetic diversity. Efficiency of breeding schemes is improved, whereas their costs could be reduced. However, the use of this new strategy may damage the quality of classical genetic evaluations.

In many countries, breeding values for dairy cattle are estimated based on an animal model using BLUP methodology. Under some hypotheses, BLUP estimates have desirable properties: they are unbiased in the sense that the expected value of the prediction is equal to the expected value of what is being predicted, and they are the best among the linear predictors in the sense that they have a minimum mean squared error of prediction (**MSE**). In the mixed model equations leading to BLUP, it has been shown that the additive genetic relationship matrix (**A**), assuming an infinitesimal model, can accommodate changes in genetic means and variances due to selection (Sorensen and Kennedy, 1983). This requires that **A** is complete and correct, that the model includes all records upon which selection is based and that pedigree is complete back to the base population.

Today, genomically enhanced breeding values (**GEBV**) are computed at the birth of candidates. It allows selection of animals not only with the highest GEBV but also with the highest Mendelian sampling contribution. This selection leads to a nonrandom set of candidates being recorded for the traits of interest. Hence, the usual assumptions on Mendelian sampling expected value and variance are no longer valid. The Mendelian sampling expected value is no longer zero so that the resulting relationship matrix is no longer the correct one. Two issues are involved. The first is selective information: data on culled animals, especially records on their daughters in the case of young sires, are missing. The second is sequential selection: selection decisions are based on GEBV, and this information is not included in the evaluation model. If the genetic evaluation model remains unchanged, when records of progeny from preselected sires are included in the model, it is feared that genetic evaluations will become biased.

The effect of genomic preselection on national and international polygenic evaluations was first discussed by the Interbull Scientific Advisory Committee (Ducrocq et al., 2008) and by van der Beek (2007). Sires' ranking, genetic parameters, and genetic trend are likely to be altered at national and international levels. The conse-

Received September 9, 2010.
Accepted October 14, 2010.
[1]Corresponding author: clotilde.patry@jouy.inra.fr

quences of using breeding values of progeny-tested bulls on genomic prediction equations were also mentioned. Due to the rapid and widespread use of genomic evaluations worldwide (11 countries were planning to start genomic evaluations in 2009–2010; Loberg and Dürr, 2009), study of the short-term consequences on the conventional evaluation system is urgently needed.

The objective of this paper was to assess via simulation the effect of a genomic preselection step on the quality of genetic evaluations. The latter is measured by the systematic estimation error (i.e., the bias), and by the accuracy of predicted genetic effects. Comparisons were made between 2 populations, one involving a genomic preselection step of young sires [genomic preselected sires (**GPS**) population] and the other involving progeny testing only (control population). Sensitivity analysis was conducted with different levels of selection intensity, heritability, and accuracy of genomic evaluation. The effect on the genetic parameter estimates was not investigated.

MATERIALS AND METHODS

Overview

Bias and accuracy of genetic evaluations were measured in the GPS and control populations. The control population was expected to provide unbiased prediction and, consequently, a higher accuracy. To demonstrate bias due to genomic preselection, breeding values (true and genomically enhanced) and performance were simulated based on a real data set coming from the Holstein population in France, which included pedigree, records, and genetic parameters of the traits under consideration. Type traits were chosen for convenience (simple model, no repeated observations).

Implementation Steps

The first key element of the chosen strategy was the identification of the candidate bulls; these consisted of the cohort of young sires (**YS**) that had only one crop of daughters in the real data set. A first BLUP evaluation based on a single trait animal model provided EBV and fixed effects estimates in the initial conditions. These EBV were used to reduce the cohort of young sires to the most promising ones on the basis of their pedigree index. Indeed, these are more likely to be genotyped by the breeding companies implementing genomic preselection. Their daughters (**D**) were then identified and their performance was deleted from the record file to mimic the data structure before a genomic preselection step. From this new data set, a second BLUP evaluation was run to obtain EBV and reliabilities (**R**), especially of young sires' parents. The EBV and R^2 were used to simultaneously simulate genomically enhanced breeding values and true breeding values (**TBV**) for the young sires and potential candidate full siblings. When genomic preselection was mimicked (GPS population), YS were selected among their full siblings, choosing the one with the highest GEBV. In both cases (with or without genomic preselection), consistent daughter records (i.e., new records) were generated, and a new BLUP evaluation was performed. As a result, EBV and TBV of the young sires and their daughters were available for bias and accuracy assessment in the 2 simulated populations. The simulation of GEBV, TBV, and daughter records, as well as the final BLUP evaluation, were repeated 50 times to measure the bias variability. Evaluations on real or simulated data were all run using an in-house BLUP software, GENEKIT, developed by the second author. The different steps of the simulation are summarized in Table 1.

Simulation of Breeding Values and Records: Principle

Let a_{YS} be the simulated TBV of a young sire and $\hat{a}_{YS}$ its EBV, solution of the animal mixed model equations. Let $\hat{a}_{YS+}$ be the GEBV of a young sire combining direct genomic value and pedigree information. A selection on this GEBV mimicked the genomic preselection of young bulls before collection of phenotypic records from daughters. If **GS** and **GD** refer to the sire and dam of a young sire YS, respectively, and σ_a^2 is the genetic variance of the trait, we have

$$a_{YS} \sim N\left(\frac{a_{GS} + a_{GD}}{2}, \sigma_a^2\right)$$
$$\hat{a}_{YS} \sim N\left(\frac{\hat{a}_{GS} + \hat{a}_{GD}}{2}, R_{YS}\sigma_a^2\right)$$
$$\hat{a}_{YS+} \sim N\left(\frac{\hat{a}_{GS} + \hat{a}_{GD}}{2}, R_{YS+}\sigma_a^2\right).$$

The GEBV reliability $\left(R_{YS+}\right)$, is a combination of the pedigree reliability $\left(R_{PED}\right)$, and the direct genomic reliability $\left(R_{GEN}\right)$, with $R_{PED} = \frac{R_{GS} + R_{GD}}{4}$. It was assumed that the genomic information contributes as much as n additional daughter records (Van Raden et al., 2009), so that $R_{GEN} = \frac{n}{n+k}$, with $k = \frac{4}{h^2} - 1$, and

Table 1. Summary of the simulation strategy and the analyses performed: description of preliminary and iterative (replicated 50 times) steps in the different populations (control, GPS)[1]

Step	Input data	Implementation	Output data of interest
Preliminary (1)	Data set 1: pedigree file + whole record file	BLUP 1	Initial EBV ($\hat{a}_0$) + fixed effects estimates (β)
Preliminary (2)	Pedigree file + $\hat{a}_0$	Identification of all the YS and their D	Target cohorts: YS (size = n) and D
Preliminary (3)	Data set 2: pedigree file + record file, cohort D removed	BLUP 2	EBV and reliabilities ($\hat{a}$, R) of the YS parents (GS, GD): $\hat{a}_{GS}, \hat{a}_{GD}, R_{GS}, R_{GD}$
Iterative (1) control	Data set 2 + $\hat{a}_{GS}, \hat{a}_{GD}, R_{GS}, R_{GD}$	Simulation of breeding values for n YS	True breeding values of the YS: a^C_{YS}
Iterative (2) control	Data set 2 + $\beta + \hat{a} + a^C_{YS}$	Simulation of D records	Data set 3: including the new records (true D records are replaced by the simulated ones)
Iterative (3) control	Data set 3	BLUP 3	$\hat{a}^C_{YS}, R^C_{YS}$
Iterative (4) control	$\hat{a}^C_{YS}, a^C_{YS}$	Bias assessment	Δ(Breeding values) = $\hat{a}^C_{YS} - a^C_{YS}$
Iterative (1) GPS	Data set 2 + $\hat{a}_{GS}, \hat{a}_{GD}, R_{GS}, R_{GD}$	Simulation of breeding values for m YS × n GS-GD pair	GEBV + TBV of the YS: $\hat{a}^{GPS}_{YS}, a^{GPS}_{YS}$
Iterative (2) GPS	$\hat{a}^{GPS}_{YS+}$	Selection of the YS based on GEBV	Cohort of genomically selected YS
Iterative (3) GPS	Data set 2 + $\beta + \hat{a} + a^{GPS}_{YS}$	Simulation of D records	Data set 3′: including the new records (true D records are replaced by the simulated ones)
Iterative (4) GPS	Data set 3′	BLUP 3′	$\hat{a}^{GPS}_{YS}, R^{GPS}_{YS}$
Iterative (5) GPS	$a^{GPS}_{YS}, \hat{a}^{GPS}_{YS}$	Bias assessment	Δ(Breeding values), comparison of reliabilities

[1]GPS = genomic preselection of young sires; YS = young sire; D = daughter; GS = sire of a YS; GD = dam of a YS; GEBV = genomically enhanced breeding value; TBV = true breeding value.

h^2 the heritability of the trait. Using Harris and Johnson's (1998) method to approximate reliability coming from different sources of information, the combined reliability $\left(R_{YS+}\right)$, was computed as

$$R_{YS+} = \frac{R_{PED} + R_{GEN} - 2R_{PED}R_{GEN}}{1 - R_{PED}R_{GEN}}.$$

In the same way, the genomically enhanced reliabilities R_{GS+} and R_{GD+} were computed at the GS and GD level, respectively. For example,

$$R_{GS+} = \frac{R_{GS} + \frac{1}{4}R_{GEN} - 2R_{GS}\frac{1}{4}R_{GEN}}{1 - R_{PED}R_{GEN}}.$$

It was assumed that each GS–GD pair has only one selected progeny. This is clearly not true, but this simplification avoids extra assumptions without inflating the importance of the son's GEBV on parental EBV. For young sires and their parents, all the breeding values (true, estimated, and genomically enhanced) may be defined together as coming from the following multivariate normal distribution:

$$\begin{bmatrix} a_{GS} \\ a_{GD} \\ a_{YS} \\ \hat{a}_{GS} \\ \hat{a}_{GD} \\ \hat{a}_{GS+} \\ \hat{a}_{GD+} \\ \hat{a}_{YS+} \end{bmatrix} \sim N\left[\begin{pmatrix} \mu_{GS} \\ \mu_{GD} \\ \mu_{YS} \\ \mu_{GS} \\ \mu_{GD} \\ \mu_{GS+} \\ \mu_{GD+} \\ \mu_{YS+} \end{pmatrix}, \begin{pmatrix} 1 & 0 & \frac{1}{2} & R_{GS} & 0 & R_{GS+} & 0 & \frac{1}{2}R_{GS+} \\ & 1 & \frac{1}{2} & 0 & R_{GD} & 0 & R_{GD+} & \frac{1}{2}R_{GD+} \\ & & 1 & \frac{1}{2}R_{GS} & \frac{1}{2}R_{GD} & \frac{1}{2}R_{GS+} & \frac{1}{2}R_{GD+} & R_{YS+} \\ & & & R_{GS} & 0 & R_{GS} & 0 & \frac{1}{2}R_{GS+} \\ & & & & R_{GD} & 0 & R_{GD} & \frac{1}{2}R_{GD+} \\ & & & & & R_{GS+} & 0 & \frac{1}{2}R_{GS+} \\ & & & & & & R_{GD+} & \frac{1}{2}R_{GD+} \\ & & & & & & & R_{YS+} \end{pmatrix}\sigma_a^2\right].$$

Using Pearson's results (Pearson, 1903) for multivariate normal distributions, GEBV and TBV were jointly simulated conditionally on the GS and GD EBV. These 2 EBV and their associated reliabilities were computed from the second BLUP evaluation on the real data set (i.e., in absence of progeny information of their sons). It follows that:

$$\begin{bmatrix} a_{GS} \\ a_{GD} \\ a_{YS} \\ \hat{a}_{GS+} \\ \hat{a}_{GD+} \\ \hat{a}_{YS+} \end{bmatrix} \sim N\left[\begin{pmatrix} \hat{a}_{GS} \\ \hat{a}_{GD} \\ \frac{\hat{a}_{GS} + \hat{a}_{GD}}{2} \\ \hat{a}_{GS} \\ \hat{a}_{GD} \\ \frac{\hat{a}_{GS} + \hat{a}_{GD}}{2} \end{pmatrix}, \mathbf{V} = \begin{pmatrix} \delta_{GS} & 0 & \frac{1}{2}\delta_{GS} & \Delta_{GS} & 0 & \frac{1}{2}\Delta_{GS} \\ 0 & \delta_{GD} & \frac{1}{2}\delta_{GD} & 0 & \Delta_{GD} & \frac{1}{2}\Delta_{GD} \\ \frac{1}{2}\delta_{GS} & \frac{1}{2}\delta_{GD} & \delta_{PED} & \frac{1}{2}\Delta_{GS} & \frac{1}{2}\Delta_{GD} & R_{YS+} - R_{PED} \\ \Delta_{GS} & 0 & \frac{1}{2}\Delta_{GS} & \Delta_{GS} & 0 & \frac{1}{2}\Delta_{GS} \\ 0 & \Delta_{GD} & \frac{1}{2}\Delta_{GD} & 0 & \Delta_{GD} & \frac{1}{2}\Delta_{GD} \\ \frac{1}{2}\Delta_{GS} & \frac{1}{2}\Delta_{GD} & R_{YS+} - R_{PED} & \frac{1}{2}\Delta_{GS} & \frac{1}{2}\Delta_{GD} & R_{YS+} - R_{PED} \end{pmatrix}\sigma_a^2\right],$$

with $\delta = 1 - R$ and $\Delta = R_+ - R$; R represents the reliability based on records and relationships information only, whereas R_+ refers to the genomically enhanced reliability. In practice, using a Cholesky decomposition of the variance-covariance matrix $\mathbf{V}$, simulated GEBV and TBV were obtained as linear combinations of 6 random standard normal variables. Consistent simulation of GEBV and TBV of full siblings was guaranteed through an

adequate repetition of the relevant random variables. Finally, the control and GPS populations differed in the way the breeding values of young sires were distributed. In the control population, random variables were drawn only once in such a way that the young sires' GEBV were normally distributed. No preselection existed. In contrast, in the GPS population, a genomic preselection step was implemented using a given selection rate defined as the proportion of retained candidates, say 1/m. In practice, TBV and GEBV were generated, on average, m times for each YS, hence mimicking genomic evaluation of m full siblings. In order to be as realistic as possible, the number of full siblings within a family i (m_i) was varied from one family to another. To ensure an expected number of m candidates with a minimum number of 1, a random number u_i was generated from a Poisson distribution $P(\lambda)$ with $\lambda = m - 1$, and the size of the full-sibling cohort was set to $u_i + 1$. Among these full siblings, the highest GEBV was assigned to the YS undergoing progeny test, whereas the remaining full siblings were culled.

Performances for the daughters of each YS were simulated in the 2 populations. From the animal model equations, these records were computed as the sum of fixed and random effects. Without loss of generality, the only fixed effect (β) considered was a contemporary group effect. In other words, data of the initial evaluation were precorrected for all nongenetic (fixed) effects and a contemporary group effect was estimated for each group of identified daughters. The breeding value of any daughter was generated as the sum of its average parental breeding value and a Mendelian sampling term (φ) drawn from a normal distribution with zero mean and a variance equal to half the genetic variance of the trait. The TBV for the sire was available as described above, whereas the TBV for the dam was generated on the basis of her EBV and reliability computed from the second BLUP evaluation when no YS progeny performances were available. A normal random variable ε was, thus, added to dam EBV to get TBV. Finally, a residual e was drawn from a normal distribution with zero mean and a variance equal to the residual variance of the trait. Let y be the record of a daughter:

$$y = \beta + \frac{a_{YS} + (\hat{a}_D + \varepsilon)}{2} + \varphi + e$$

with

$$\left\{\begin{array}{l} \varepsilon \sim N[0,(1-R_D)\sigma_a^2] \\ \varphi \sim N(0,\frac{1}{2}\sigma_a^2) \\ e \sim N(0,\sigma_e^2) \end{array}\right\}.$$

Numerical Applications

The simulations were first done following the national data set for a conformation trait, udder depth (**UD**), for the Holstein breed in France. A total of 4,110,229 records were available, and the pedigree file included 5,917,739 animals. The young sires were chosen among those born in 2001, 2002, and 2003, having more than 10 and fewer than 150 recorded daughters. In all, 1,875 sires fulfilled these criteria. Among them, 799 young sires were selected with their 40,222 daughters. Selection bias was first assessed for the UD trait with a heritability of 0.36, assuming that the top 25% (proportion, $\mathbf{p} = 0.25$) of the young sires were selected after genomic evaluation. The GEBV simulations also required choosing the number of daughters that would provide the same increase in reliability as that of the genomic evaluation. In the North American Holstein population, Van Raden et al. (2009) reported 9 daughter equivalents from genomic prediction for body depth with a heritability of 0.37. Here, an initial value of 10 genomic equivalent daughter contributions (**gEDC**) was chosen, which is equivalent to a reliability of genomic evaluation of 50%. This set of parameters ($gEDC = 10$, p = 0.25, $h^2 = 0.36$) defines the udder depth trait reference scenario, **UD_REF**.

Sensitivity Analysis

In order to assess the magnitude of the bias due to genomic preselection and to understand the role of parameters such as the heritability of the trait, intensity of genomic selection, and reliability of genomic evaluations, various scenarios were implemented. Compared with that of the UD_REF scenario, the value of heritability was decreased ($h^2 = 0.14$): simulations were performed for another conformation trait, foot angle (**FA**). Notice that decreasing heritability makes the genomic evaluation less accurate if gEDC is kept equal to 10. The gEDC was increased to 26, which is possibly optimistic but made to maintain the same level of genomic accuracy. Even then, the cohorts of YS were no longer the same for FA. To keep fairly comparable scenarios, whereas handled traits are different, the YS for the FA trait were chosen so that the same average pedigree index was observed in both situations. A total of 601 young sires were, thus, identified with their 31,976 daughters. It defined the scenario hereafter called the foot angle trait reference scenario, **FA_REF**. Then the proportion of selected animals was decreased from 0.25 to 0.10 for both traits, defining scenarios udder depth trait and foot angle trait, each with proportion of selected animals at 0.10 (**UD_p** and **FA_p**, respectively). Finally, for the FA trait, 2

Table 2. Parameters for the different studied scenarios

Scenario[1]	Proportion of selected candidates	Heritability	Genomic equivalent daughter contribution	Young sires (n)
UD_REF	0.25	0.36	10	799
UD_p	0.10	0.36	10	799
FA_REF	0.25	0.14	26	601
FA_p	0.10	0.14	26	601
FA_gEDC	0.25	0.14	10	601

[1]UD = udder depth trait; FA = foot angle trait; REF = reference; p = proportion of young sires retained; gEDC = genomic equivalent daughter contribution.

levels of genomic accuracy were compared. In the scenario called **FA_gEDC**, gEDC was decreased from 26 (in FA_REF) to 10. Reliability of direct genomic values decreased from 50% to 27%. In order to sum up (see Table 2), UD_REF and FA_REF were compared to assess the effect of heritability change. Having UD_REF and UD_p on one side, and FA_REF and FA_p on the other side, highlighted the effect of selection intensity. The comparison of FA_REF with FA_gEDC showed the importance of genomic accuracy.

Criteria for Assessing the Quality of the Genetic Evaluations

Bias was measured as the difference between TBV and EBV. Many criteria exist to assess the precision of evaluations; we retained 2 of them. The mean square error of prediction was preferred to the prediction error variance, as used in other studies on selection processes (Sorensen and Kennedy, 1984; Van Vleck, 1987; Schenkel et al., 2002). Hickey et al. (2008) and L. R. Schaeffer (Department of Animal and Poultry Science, University of Guelph, Canada, personal communication) considered the computation of the correlation between true and EBV. We decided to use the squared correlation to compare it to the approximate reliability, a byproduct of the mixed model equations. This reliability is computed using the standard deviation in the base population. Statistics were averaged over 50 replicates. Definitions of all criteria assessing the quality of genetic evaluations are summed up in Table 3.

RESULTS

The results are first presented for the reference scenario (UD_REF). To facilitate the interpretation, breeding values are expressed in genetic standard deviation of the trait.

Validation of the Approach

When the selection scheme involves progeny testing, information is collected for the whole population of selection candidates so that the estimation of the Mendelian sampling term is consistent with the hypotheses underlying the BLUP methodology. In particular, the expected value of the Mendelian sampling term is supposed to be zero. For both populations simulated in the UD_REF scenario, Table 4 displays genetic values (TBV, EBV, MS) and the difference between EBV and TBV. In the control population, EBV are not significantly different from TBV and the MS term is not significantly different from 0. Hence, the considered population structure and estimation method lead to unbiased evaluations in classical conditions. And, in the GPS population, TBV are indeed significantly larger than the TBV observed in the control population (Table 4), illustrating the fact that genomic preselection was effective.

Table 3. Definitions of the criteria for assessing quality of genetic evaluations with replicate r and $a_r = \overline{TBV}$, $\hat{a}_r = \overline{EBV}$ [1]

Bias (Δ)	Mean square error (MSE)	Squared correlation (ρ^2)
$= \frac{1}{50}\sum_{r=1}^{50} E(\hat{a}_r - a_r)$	$= \frac{1}{50}\sum_{r=1}^{50}\left[Var(\hat{a}_r - a_r) + E(\hat{a}_r - a_r)^2\right]$	$= \left[\frac{1}{50}\sum_{r=1}^{50}\frac{\mathrm{cov}(\hat{a}_r, a_r)}{\sqrt{\mathrm{var}(\hat{a}_r)\,\mathrm{var}(a_r)}}\right]^2$

[1]TBV = true breeding value; E = expected value; Var = variance.

Table 4. Mean (first line) and standard deviation within replicate (second row, in italics) of genetic values and bias (Δ) in the cohort of selection candidates when genomic selection is implemented (GPS) or not (control)[1]

Population	TBV	EBV	MS	$\Delta = \overline{EBV - TBV}$
Control	0.931	0.933	−0.001 (NS)	0.002 (NS)
	0.864	*0.742*	*0.630*	*0.427*
GPS	1.384	1.238	0.304***	−0.146 ***
	0.781	*0.694*	*0.578*	*0.409*

[1]TBV = true breeding value; MS = Mendelian sampling contribution. Values are averaged over 50 replicates. Tested the null hypothesis H_0, where the mean of MS or mean of bias equal zero.
***$P < 0.001$.

Bias Evidence

Table 4 indicates that the MS term in the GPS population is significantly larger than zero, whereas the variance has decreased compared with that in the control population. From this first scenario, the basic hypothesis on the MS distribution is shown to be no longer consistent with the observed distribution. Finally, the difference between EBV and TBV is significant and negative ($\Delta = -0.146$), indicating that evaluations are biased when genomic preselection is implemented. Furthermore, the expected value of the bias (Δ) between the 50 replicates presents a low variability in the control (standard deviation, $\sigma_\Delta = 0.016$) and in the GPS ($\sigma_\Delta = 0.014$) populations, showing that the results on bias are not random.

Distribution of the Bias in Genetic Evaluations

Under the classical selection scheme, the difference between EBV and TBV is distributed around a mean value of 0 with a standard deviation within replicate of 0.427, as indicated in Tables 4 and 5. This is a symmetric distribution where mean and median are superposed, the distance between third and second, and second and first quartiles are the same (Table 5). It displays the distribution of the prediction error which is minimized with BLUP. When genomic preselection is implemented, such a type of distribution is also observed, but values are systematically translated into more negative values. On average, bias is equal to −0.146 with a standard deviation between replicates of 0.014 in the YS cohort and bias is equal to −0.044 with a standard deviation between replicates of 0.006 in the cohort of their daughters. Hence, EBV underestimate TBV.

Effect on the Accuracy of Evaluations

By construction, the BLUP solutions should maximize the correlation between TBV and EBV and minimize the mean square error of prediction. The MSE increased in the GPS population compared with those of the control population. Similarly, the squared correlation between TBV and EBV was lower in the GPS population than that in the control population (Table 6), whereas the amount of information was identical, as indicated by the value of the approximate reliability (REL). All of the indicators show that the genomic preselection step decreased the accuracy of the estimations compared with that of the control population, with a diluted effect on daughters compared with that of young sires.

Sensitivity Analysis

Table 7 indicates that all parameters (intensity of selection, heritability, and accuracy of genomic evaluation) induced changes in the mean bias and mean MSE

Table 5. Quartiles and first moments of the difference between EBV and true breeding values for the young sires (YS) and their daughters (D) when genomic selection is implemented (GPS) or not (control)[1]

	Control population		GPS population	
Item	YS	D	YS	D
Minimum	−2.249	−3.523	−2.598	−3.655
First quartile	−0.272	−0.486	−0.406	−0.539
Median	0.004	0.011	−0.138	−0.043
Mean	0.002 (NS)	0.011**	−0.146***	−0.044***
SD between replicates	0.016	0.006	0.014	0.006
Third quartile	0.278	0.508	0.127	0.454
Maximum	2.195	3.700	1.915	3.608
SD within replicate	0.427	0.737	0.409	0.736

[1] Tested the null hypothesis H_0, where the mean of MS or mean of bias equal zero.**$P < 0.01$; ***$P < 0.001$.

Table 6. Mean squared error (MSE) and reliability measures [ρ^2(EBV, TBV), REL] in control and GPS populations for udder depth trait and 25% selection rate[1]

Population	MSE	ρ^2 (EBV, TBV)	REL
Control			
Young sires	0.183	0.756	0.815
Daughters	0.544	0.414	0.476
GPS			
Young sires	0.188	0.727	0.815
Daughters	0.544	0.394	0.476

[1]ρ^2 = squared correlation; TBV = true breeding value; REL = approximate reliability; GPS = genomic preselected sires.

of the BLUP evaluations compared with those of the reference scenario, UD_REF. When genomic selection intensity increased, the magnitude of the bias also rose and the accuracy of the estimations decreased. This is verified for UD and FA traits when the proportion of selected candidates decreased from 25 to 10% (scenario UD_REF compared with UD_p and scenario FA_REF compared with FA_p). The same trends were observed when the precision of the genomic evaluation decreased (FA_gEDC compared with FA_REF). Finally, scenarios UD_REF and FA_REF were compared, with the same proportion of selected candidates and the same amount of genomic information. When the heritability was lower, genetic evaluations were more affected: the magnitude of the bias was larger and the accuracy of the estimations was lower.

DISCUSSION

Simulating data such as TBV and phenotypic values has many advantages. Compared with studies on real data, the only relevant source of bias is isolated and the mechanisms underlying bias are controlled so that it is then easy to vary parameters of interest and better understand mechanisms underlying bias. In the literature, several methods were proposed to systematically detect bias in prediction procedures. They involved assessing the change between estimates of 2 sets of selected and unselected data to study sequential selection (Mallinckrodt et al., 1995) or between 2 consecutive predictions (Reverter et al., 1994). However, by accessing true genetic values, the actual properties of BLUP solutions may be contrasted with their expectation of unbiasedness and of minimum error variance. The major limitation with this method is that it cannot be used to check routinely if classical evaluations are biased by genomic preselection.

The approach in our study was designed to be realistic. It is based on real data with large numbers of herds and young sires promoting genetic and environmental diversity. Moreover, the choice of paternal families and their composition (number of siblings) tried to mimic the strategies implemented by breeding companies when genotyping candidates. However, only one data structure was considered with type traits. Repeated or longitudinal data, for example, for which evaluation models are different, were not investigated.

Genomic selection can be seen as a missing data process where only progeny of selected animals are recorded. These were selected based on GEBV; however, this information was not included in the evaluation. Actually, whatever the scenario, it is always observed that under genomic selection, bias is significant and breeding values of selected young sires and their daughters are underestimated. Mixed model equations assume that selected young sires have a Mendelian sampling term with zero expectation, whereas the MS contribution is now, on average, positive. The accuracy of evaluations is also always lower when a genomic preselection step is implemented. The magnitude of the bias and the accuracy loss are even greater when the intensity of genomic selection is increasing. This is not surprising, because the proportion of missing data is indeed rising. We observe the same trends when the accuracy of the genomic evaluations is decreasing, when the correlation between GEBV and the trait is in the same way declin-

Table 7. Mean bias (Δ), standard deviation (σ_Δ), and mean squared error (MSE) over 50 replicates when the heritability, the proportion of selected candidates, and the reliability of genomic evaluation vary

Scenario[1]	Δ,σ_Δ	MSE
UD_REF	−0.146 ± 0.014	0.188
UD_p	−0.227 ± 0.016	0.217
FA_REF	−0.214 ± 0.021	0.305
FA_p	−0.155 ± 0.020	0.299
FA_gEDC	−0.249 ± 0.025	0.364

[1]UD = udder depth trait; FA = foot angle trait; REF = reference; p = proportion of young sires retained; gEDC = genomic equivalent daughter contribution; UD_REF = 36% of heritability, 25% of young sires retained, 10 EDC from genomic evaluation; UD_p = 36% of heritability, 10% of young sires retained, 10 EDC from genomic evaluation; FA_REF = 14% of heritability, 25% of young sires retained, 26 EDC from genomic evaluation; FA_p = 14% of heritability, 10% of young sires retained, 26 EDC from genomic evaluation; FA_gEDC = 14% of heritability, 10% of young sires retained, 10 EDC from genomic evaluation.

ing. Finally, Sorensen and Kennedy (1984) reported that the magnitude of a selection bias also depended on heritability. The coefficients of the relationship matrix being incorrect and even combined with the variance ratio in the mixed model equations may explain how bias and MSE vary with heritability.

The magnitude of the bias was significantly different from zero under genomic selection in all of the studied scenarios. Over one generation only, the bias comprises between 4 and 11% of genetic standard deviation for the daughters of the young sires, and between 15 and 25% for the young sires. For the same range of heritability and selection intensity, such magnitude was never reported in studies assessing bias in genetic evaluations. As genomic evaluations are predicted to be the future dominant selection tool (progeny testing is already being abandoned in some breeding programs), the preselection intensity, which is not yet accounted for in classical models, may become large, and clearly biased evaluations are to be feared. Henderson (1990a, b) reported that a biased predictor may exist that has a smaller mean squared error than an unbiased predictor. Indeed, Gianola et al. (1988) and L. R. Schaeffer (Department of Animal and Poultry Science, University of Guelph, Canada, personal communication) used this argument to stress that the most important aspect to increase genetic gain is the increase of EBV accuracy. However, our results indicate an increase of the MSE. Assuming genomic preselection of the young sires only, genetic progress may then be decreased by 3 to 4%, if 25 or 10% of YS are retained. It means that an uncontrolled loss of accuracy can actually threaten the effectiveness of a breeding program and have large genetic and then economic effects. It seems essential to take into account this bias. To assess the bias magnitude over more generations, a fully stochastic study would have to be used. Considering previous studies, we expect that effect on genetic evaluations would increase. For example, Schenkel et al. (2002) found that the bias increased over generations in the case of missing pedigree information. In fact, direction and amplitude of the bias may particularly become unpredictable when daughters of selected young sires are mated to other preselected young sires.

CONCLUSIONS

This study presents evidence that national genetic evaluations are biased once young sires are preselected based on genomic breeding values. The challenge is now to account for genomic preselection in the classical evaluation models. Understanding why evaluations are biased leads us to propose the inclusion of, in the classical evaluation, performance obtained from deregressed GEBV of culled and selected candidates, whereas Misztal et al. (2009) proposed a single-step genetic evaluation combining pedigree-based information with genomic data. All information upon which selection has been based would then be included. In addition, genomic and polygenic information would be naturally combined so that the level of accuracy of national evaluations would be increased. However, a correction, at a national level, would be even more relevant for international evaluations where breeding practices strongly differ between countries. At least, a validation test would be required to check that national evaluations are still unbiased despite genomic preselection to ensure the quality of international evaluations.

ACKNOWLEDGMENTS

Financing from Agence Nationale de la Recherche project AMASGEN (Jouy-en-Josas, France) and APISGENE are acknowledged. The authors thank members of the AMASGEN project for their comments: P. Croiseau, S. Fritz, F. Guillaume, A. Legarra, C. Robert-Granie.

REFERENCES

Ducrocq, V., M. Goddard, L. Schaeffer, and G. Banos. 2008. Interbull Scientific Advisory Committee Annual Report (2007–2008) to the Interbull Steering Committee. 4 pp. Interbull meeting, Niagara Falls, NY.

Gianola, D., S. Im, and R. L. Fernando. 1988. Prediction of breeding value under Henderson's selection model: A revisitation. J. Dairy Sci. 71:2790–2798.

Harris, B., and D. Johnson. 1998. Approximate reliability of genetic evaluations under an animal model. J. Dairy Sci. 81:2723–2728.

Henderson, C. R. 1990a. Accounting for selection and mating biases in genetic evaluation. Pages 413–436 in Advances in Statistical Methods for Genetic Improvement of Livestock. D. Gianola and K. Hammond, ed. Springer-Verlag, Heidelberg, Germany.

Henderson, C. R. 1990b. Statistical methods in animal improvement: Historical overview. Pages 2–14 in Advances in Statistical Methods for Genetic Improvement of Livestock. D. Gianola and K. Hammond, ed. Springer-Verlag, Heidelberg, Germany.

Hickey, J. M., M.G. Keane, D. A. Kenny, A. R. Cromie, H. A. Mulder, and R. F. Veerkamp. 2008. Estimation of accuracy and bias in genetic evaluations with genetic groups using sampling. J. Anim. Sci. 86:1047–1056.

Loberg, A., and J. Dürr. 2009. Interbull survey on the use of genomic information. Pages 3–14 in Proc. Interbull International Workshop on Genomic Information in Genetic Evaluations. Interbull Bulletin No. 39. Interbull, Uppsala, Sweden.

Mallinckrodt, C. H., B. L. Golden, and R.M. Bourdon. 1995. The effect of selective reporting on estimates of weaning weight parameters in beef cattle. J. Anim. Sci. 73:1264–1270.

Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. J. Dairy Sci. 92:4648–4655.

Pearson, K. 1903. Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. Philos. Trans. Royal Soc. London 200:1–66.

Reverter, A., B. L. Golden, R. M. Bourdon, and J. S. Brinks. 1994. Technical note: Detection of bias in genetic predictions. J. Anim. Sci. 72:34–37.

Schenkel, F. S., L. R. Schaeffer, and P. J. Boettcher. 2002. Comparison between estimation of breeding values and fixed effects using Bayesian and empirical BLUP estimation under selection on parents and missing pedigree information. Genet. Sel. Evol. 34:41–59.

Sorensen, D. A., and B. W. Kennedy. 1983. The use of the relationship matrix to account for genetic drift variance in the analysis of genetic experiments. Theor. Appl. Genet. 66:217–220.

Sorensen, D. A., and B. W. Kennedy. 1984. Estimation of response to selection using least-squares and mixed model methodology. J. Anim. Sci. 58:1097–1106.

van der Beek, S. 2007. Effect of genomic selection on national and international genetic evaluations. Pages 115–118 in Proc. Interbull Meeting: Dublin, Ireland. Interbull Bulletin No. 37. Interbull, Uppsala, Sweden.

VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. *Invited review:* Reliability of genomic predictions for North American Holstein bulls. J. Dairy Sci. 92:16–24.

Van Vleck, L. D. 1987. Contemporary groups for genetic evaluations. J. Dairy Sci. 70:2456–2464.

Erratum – article I

1) Table 3 should be replaced by the following, the definition of the squared correlation being modified:

Definitions of the criteria for assessing quality of genetic evaluations with replicate r and $a_r = \overline{TBV}, \hat{a}_r = \overline{EBV}$

Bias	MSE	Squared correlation
$\Delta = \frac{1}{r}\sum_{r=1}^{50} E(\hat{a}_r - a_r)$	$MSE = \frac{1}{r}\sum_{r=1}^{50}(Var(\hat{a}_r - a_r) + E(\hat{a}_r - a_r)^2)$	$\rho^2 = \frac{1}{r}\sum_{r=1}^{50}(\frac{\text{cov}(\hat{a}_r, a_r)}{\sqrt{\text{var}(\hat{a}_r)\,\text{var}(a_r)}})^2$

2) Table 7 should be replaced by the following, results for scenario Fa_p being modified:

Mean bias (Δ*), standard deviation (*σ_Δ*), and mean squared error (MSE) over 50 replicates when the heritability, the proportion of selected candidates and the reliability of genomic evaluation vary*

Scenarios	Δ, σ_Δ	MSE
UD_REF	-0.146 +/- 0.014	0.188
UD_p	-0.227 +/- 0.016	0.217
FA_REF	-0.214 +/- 0.021	0.305
FA_p	-0.338 +/- 0.020	0.364
FA_gEDC	-0.249 +/- 0.025	0.339

CHAPTER 5 - Accounting for a genomic selection step in national evaluations

The objective of the present chapter is to highlight the necessity of accounting for genomic selection in genetic evaluations. Possible methods to avoid this bias will be addressed. One strategy is implemented at the national level in Patry and Ducrocq (2011a).

5.1. Necessity to account for bias

5-1.1. LONG TERMS PROBLEMS DUE TO BIAS IN GENETIC EVALUATIONS

Bias as a systematic under- or overestimation of breeding values may reduce response to selection (Vinson, 1987).

In fact, any bias in breeding value estimation generates a number of problems regarding accuracy and genetic parameter estimates which decrease the quality of genetic evaluations to finally slow down genetic gain.

Indeed, the reliability of genetic evaluations is by definition the squared correlation between true and estimated breeding values. Consequently, biased estimates for a fraction of the population are directly associated with a loss of precision.

Furthermore, any bias in evaluations may cumulate over generations and as a result, the estimation of genetic parameters, i.e., variances, heritability, and genetic correlation between traits or countries, may be affected too. Mallinckrodt *et al.* (1995) reported a reduction in correlation between traits because of selective reporting.

International rankings are especially sensitive to over- or underestimation of national genetic trends (Boichard *et al.*, 1995, Ducrocq *et al.*, 2003). If any, national or international selection decisions will deviate from optimality. This will reduce the efficiency of breeding programs.

5-1.2. BIAS DUE TO GENOMIC SELECTION: A LARGE-SCALE AND URGENT PROBLEM

Regarding the first simulation results, BLUP solutions of the young bulls were systematically underestimated after genomic selection. It is now relevant to assess if this issue is important from a practical point of view.

- **Why is genomic selection more worrying than marker-assisted selection?**

Marker-assisted selection has been rarely implemented in the world, except in France where MAS programs were run from 2001 to 2009. It was first used to apply a within family selection using microsatellites markers in the Holstein, Normande and Montbéliarde populations. About 60,000 animals were genotyped in these programs. However, animals had

only reduced molecular information and the Mendelian sampling deviation from the null expectation was small compared to what is expected today with genomic selection, so the risk of significant bias was ignored and probably inexistent.

Given the genome-wide dense information available, the relatively high accuracy of MS term estimation, the rather strong selection intensity which is applied with genomic selection, and mainly the disuse of progeny testing, the concern for bias due to genomic selection not accounted for in the evaluation models is now clearly justified.

- **The fast and wide adoption of genomic selection**

In less than 4 years, 16 countries have developed genomic evaluations on 7 breeds or groups of breeds. An international genomic evaluation is under development (Van Raden and Sullivan, 2010) and would also benefit to countries without a genomic evaluation system, broadening even more the use of genomic tools worldwide. The implementation of genomic selection with the modification of breeding schemes is occurring very fast and in such a wide way that it cannot be neglected.

- **An urgent need to account for genomic selection in classical evaluations**

To illustrate this issue and understand how genomic selection may quickly impact the classical evaluation, the case of France can be examined. France was one of the first countries to implement genomic selection. In June 2009, GEBV were officially published for the first time. These selected bulls were already sexually mature at that time and their first daughters were born in March 2010.

Considering the tightest possible time frame, these first daughters were bred 15 months later, i.e., in May 2011. Hence, some daughters of the first generation of genomically selected sires will calve and begin their first lactation in March 2012. Official evaluations occur 3 times a year, in February, June and October. The first production records are usually included 90 days after the beginning of the lactation. For these cows, records will be analyzed in June or October 2012. Therefore, daughter performances on complete lactation for the first genomically selected sires may be included in the official evaluation as soon as February 2013.

Conclusion: Any bias in BLUP solutions and loss in precision is to some extent harmful for selection efficiency. Violations of BLUP assumptions after a genomic selection step are obvious. The use of genomic information is fast and widespread across the world. The consequences will fast be reality: accounting for genomic selection in national evaluation models is especially urgent.

5.2. Approaches to account for selection

5-2.1. USUAL APPROACHES TO ACCOUNT FOR BIAS IN BLUP EVALUATIONS

In the history of livestock improvement, BLUP-based methodologies to estimate breeding values have continuously been improved to avoid bias. All approaches rely on the general principle that all available information to describe the selection process or the distribution of the selection criteria should be included in one way or another. Specific treatments were developed to face the identified sources of bias. They are listed in Table 2.

Table 2: Accounting for selection bias or other sources of bias in BLUP evaluations

Sources of bias	Treatments/recommendations
Parental selection or assortative mating	Include information on dams and account for their genetic merit to make the genetic relationship matrix complete and correct (Kennedy *et al.*, 1988)
Sequential selection	Implement multi-trait evaluations (Pollak *et al.*, 1984)
Missing pedigree information	Define genetic groups for unknown parents of the base population (Quaas, 1988)
Use of highly selected animals from foreign countries	Define groups of unknown parents according to country of origin to account for different genetic levels between base population in different countries. Then, blend information: include foreign information (foreign EBV) into national genetic evaluations to account for the genetic merit of these foreign sires for example. EBV can be de-regressed and transformed into average records of "virtual daughters", the number of "virtual daughters" being derived from the sire reliability (Bonaiti and Boichard, 1995, Pedersen *et al.*, 1999)
Heterogeneity of variance among subclasses	At least, 4 alternatives: - Define contemporary groups to account for similar conditions and management practices; - Adjust variances across subclasses (e.g., by log transformation or by scaling of observations to a constant standard deviation); - Implement multi-trait evaluations; - Use a model where logarithm of residual variances are described as a function of fixed (and possibly random) effects (Foulley *et al.*, 1990).
Preferential treatment (PT)	Collect or report all necessary information and define new fixed effects (for example, use of growth hormone, special herd management group)

From past experiences on bias in BLUP solutions and especially on bias due to selection, several non exclusive alternatives can be considered to account for genomic selection:

- Include all animals, i.e., the culled and the selected candidates in a unique genetic relationship matrix so they are all considered in the unique analysis;
- Blend information: GEBV could be de-regressed and weighed by their reliability to mimic a certain number of "virtual daughters" records, to be combined with information from actual phenotypes;
- Implement a multi-trait evaluation and consider GEBV as a correlated trait of EBV on which selection was based;
- Define a new fixed effect such as a contemporary group effect for contemporary young bulls undergoing genomic selection and to account for their difference in genetic levels.

The general principle is to include all genotyped candidates and the selection criteria on which genomic selection is based or any element describing the selection process.

5-2.2. ALTERNATIVE APPROACHES TO ACCOUNT FOR GENOMIC SELECTION

- **A necessity: combining genomic with polygenic information**

Selected candidates have both molecular marker information and classical information, i.e. phenotypic and pedigree-based information. Culled candidates have only molecular marker information. Consequently, there is a need to combine both types of information to analyze together all animals in the same analysis. Four advantages can be listed for such a strategy:

- it would transfer genomic information from genotyped animals to their relatives which are not genotyped;
- conversely, performances on phenotyped relatives would also enrich the genomic evaluations of the young bulls;
- it would account for selective genotyping and prevent from bias in genomic predictions;
- it would account for genomic selection and prevent from bias in classical evaluations.

There are two alternative approaches to combine information in genetic evaluations. In chapter 2 (Section 2-1.3), we reviewed various methods to combine all types of information in genetic evaluations. On the one hand, the more natural approach would be the single step one but it is not yet routinely implemented. On the other hand, multi-step approaches are routinely implemented and provide GEBV to the culled genotyped candidates. The challenge is to include these GEBV in classical evaluations for all genotyped candidates.

- **A genomic pseudo-performances approach (Ducrocq and Liu, 2009)**

At the Interbull congress in Barcelona (Spain), Ducrocq and Liu (2009) proposed a method to include genomic pseudo-performances for genotyped animals in HMME. The principle is to

de-regress the GEBV available for all genotyped individuals and to include these de-regressed values as if candidates had own daughters with records. These pseudo-performances require to be weighted in the HMME to reflect the actual amount of information coming from the genomic information. Van Raden (2009) proposed to convert the direct genomic reliability R^2_{DGV} from genomic predictions into daughter equivalents. Here we define genomic EDC from genomic contribution applied to animal model:

$$gEDC = \frac{kR^2_{DGV}}{1-R^2_{DGV}} \text{[25] with } k = \frac{4-h^2}{h^2}$$

However, Ducrocq and Liu (2009) proposed two other alternatives to compute gEDC: they distinguished two cases, when genomic predictions were obtained either from G-BLUP or after estimation of marker effects from a training population. These alternatives will be described further in details in chapter 7.

5.3. Our implemented strategy

5-3.1. Choices for de-regression method and gEDC computation

The method of Ducrocq and Liu (2009) was implemented on real data using the same simulation framework as in the first study on bias assessment (Patry and Ducrocq, 2011b). The methods for de-regression and for computation of gEDC had to be chosen.

De-regression is a common practice in genetic evaluations. There are two types of procedures: 1) back solving the HMME to remove covariance between related animals (Jairath *et al.*, 1998) as implemented as a first step in international genetic evaluations: all animals are considered at a time or 2) individually standardizing proofs by adjusting for a mean (i.e., the average breeding value or average parent average of the population of interest) and dividing by a standard deviation based on reliability. In our study, GEBV were simulated and could be de-regressed using the first approach.

It was decided to begin with the simplest case for gEDC. The same value was computed for all genotyped candidates based on the genomic prediction reliability. Various levels of genomic reliability were converted into gEDC: it could be interpreted as various amounts of genomic information.

5-3.2. Alternative scenarios

The two main scenarios of the previous study on bias assessment, i.e. 1) and 2), were compared to two new scenarios derived from 2) and including pseudo-performances.

- Scenario 1: All candidates were included in HMME after progeny testing of their daughters: they all had actual records (Control scenario);

- Scenario 2: Only the selected candidates had actual records and were included in HMME;
- Scenario 3: Genomic pseudo-performances were included only for the selected genotyped candidates;
- Scenario 4: Genomic pseudo-performances were included for all genotyped candidates.

5-3.3. EFFICIENCY OF THE METHOD IN REDUCING BIAS MAGNITUDE

Bias was measured as the difference between true and estimated breeding values in the four scenarios. Bias was measured among the cohort of young sires undergoing genomic selection and the cohort of their daughters. It was expected that this difference would be zero in the first scenario and close to zero in any scenario accounting for genomic selection in a satisfying way. Reliability and MSE were also measured in each scenario to assess the quality of the classical evaluations.

This strategy was implemented: methods and results were presented in a second article: Patry, C. and V. Ducrocq. 2011. Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. Genet Sel Evol 43:30.

5-3.4. MAIN CONCLUSIONS OF THE SIMULATION STUDY

Simulations showed that bias was removed from BLUP breeding values only when all information about selection candidates was included, i.e., in scenario 4. The method performed well whatever the selection intensity. But it performed better with a higher heritability trait and when reliability of genomic evaluation was high enough.

This approach is promising: conceptually it is possible to account for genomic selection in national evaluations and avoid biased breeding values. Based on a multi-step approach, it might be easily and quickly implemented although some adjustments are necessary to properly compute gEDC

.

Article II

Patry, C. and V. Ducrocq. 2011.

Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle.
Genet Sel Evol 43:30.

Patry and Ducrocq *Genetics Selection Evolution* 2011, **43**:30
http://www.gsejournal.org/content/43/1/30

GSE Genetics Selection Evolution

RESEARCH Open Access

Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle

Clotilde Patry[1,2*] and Vincent Ducrocq[1]

Abstract

Background: In future Best Linear Unbiased Prediction (BLUP) evaluations of dairy cattle, genomic selection of young sires will cause evaluation biases and loss of accuracy once the selected ones get progeny.

Methods: To avoid such bias in the estimation of breeding values, we propose to include information on all genotyped bulls, including the culled ones, in BLUP evaluations. Estimated breeding values based on genomic information were converted into genomic pseudo-performances and then analyzed simultaneously with actual performances. Using simulations based on actual data from the French Holstein population, bias and accuracy of BLUP evaluations were computed for young sires undergoing progeny testing or genomic pre-selection. For bulls pre-selected based on their genomic profile, three different types of information can be included in the BLUP evaluations: (1) data from pre-selected genotyped candidate bulls with actual performances on their daughters, (2) data from bulls with both actual and genomic pseudo-performances, or (3) data from all the genotyped candidates with genomic pseudo-performances. The effects of different levels of heritability, genomic pre-selection intensity and accuracy of genomic evaluation were considered.

Results: Including information from all the genotyped candidates, i.e. genomic pseudo-performances for both selected and culled candidates, removed bias from genetic evaluation and increased accuracy. This approach was effective regardless of the magnitude of the initial bias and as long as the accuracy of the genomic evaluations was sufficiently high.

Conclusions: The proposed method can be easily and quickly implemented in BLUP evaluations at the national level, although some improvement is necessary to more accurately propagate genomic information from genotyped to non-genotyped animals. In addition, it is a convenient method to combine direct genomic, phenotypic and pedigree-based information in a multiple-step procedure.

Background

In dairy cattle, selection decisions on candidates are now widely based on Genomically Enhanced Breeding Values (GEBV) instead of Estimated Breeding Values (EBV) obtained after progeny testing. Together with the increasing availability of genotypes, further methodological developments are expected to increase the reliability of GEBV and to achieve higher genetic progress.

One challenge is to combine genomic and non-genomic information for all the animals, whether they are genotyped or not. Indeed, the number of genotyped animals is still small compared to the number of non-genotyped animals with phenotypes. Having animals with both EBV and GEBV and other animals with EBV only creates some uncertainty for breeding companies and farmers on how to optimally choose among the candidates for selection. It is also desirable to use all available information, whether genomic, phenotypic or pedigree-based, to assess the additive genetic value of any animal. Currently, there are two alternative procedures to combine data, either a multi-step procedure [1,2], which is based on selection index theory, or a single-step procedure (SSP) based on a relationship matrix that blends full pedigree and genomic information to simultaneously evaluate genotyped and non-genotyped animals [3-5]. How to correctly propagate information from genotyped to non-genotyped animals without overestimating reliabilities and without biasing breeding values remains an issue [4,6].

* Correspondence: clotilde.patry@jouy.inra.fr
[1]INRA, UMR 1313 Génétique Animale et Biologie Intégrative, F-78350 Jouy-en-Josas, France
Full list of author information is available at the end of the article

BioMed Central

© 2011 Patry and Ducrocq; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Including genotyped and non-genotyped animals in a single genetic analysis is also necessary to properly account for biases due to selective genotyping [7] or phenotyping [4,6,8]. The latter corresponds, for example, to young sires that are pre-selected based on genomic information: only sires with higher GEBV and hence with a higher Mendelian sampling term receive phenotypes from daughters a few years after pre-selection. BLUP (Best Linear Unbiased Prediction) assumes that Mendelian sampling terms have zero expectation [9]. Thus, genomic pre-selection (GPS) leads to biased EBV and reduced accuracy in national genetic evaluations based on a polygenic model [10]. In France, genomic evaluations became official in 2009. Since then, bulls that were pre-selected according to genomic information have been used. In 2013, the first records of their daughters will be included in the national BLUP evaluation and the resulting EBV might be biased. One concern is that biased EBV and their corresponding daughter yield deviations (DYD) may impact the estimation of SNP effects in subsequent years. This issue is also relevant at the international level, since the trade of bull semen is based on EBV from Multiple Across Country Evaluations (MACE) that are computed assuming unbiased national EBV. With genomic pre-selection more and more widely implemented, accounting for such practices is becoming very important.

Ducrocq and Liu [6] proposed a method to include genomic information in national BLUP evaluations. The approach consists of de-regressing all GEBV on which pre-selection was based, using the effective contribution of the additional genomic information as the weight. Then, all the genotyped candidates receive a pseudo-record based on genomic information to be included in the mixed model equations (MME), in addition to the actual phenotypic records. The BLUP model assumption that all sources of information on which selection is based are included is then fulfilled.

The aim of this study was to implement such a method and to assess its ability to remove bias due to genomic pre-selection of young sires. In the study of Patry and Ducrocq [10], actual data were used to simulate breeding values and mimic genomic pre-selection of the last generation of sires to assess bias in national BLUP evaluations. In the current study, the same population and simulated data as in [10] were used to measure bias before and after including genomic information. In addition, the issue of combining genomic with traditional information, i.e. phenotypes and pedigree, is addressed.

Methods

Overview

Data were generated as described in Patry and Ducrocq [10] and GEBV were simulated for a cohort of young sires that was considered as a cohort of selection candidates. GEBV were used to retain a proportion of the best candidates, mimicking genomic pre-selection. To account for this selection step in BLUP evaluations at the national level, GEBV were de-regressed to provide genomic pseudo-performances for all the genotyped candidates. A weight derived from the increase in reliability of EBV due to genotype information was associated to each pseudo-performance. Pseudo-performances and their associated weights were included in Henderson's mixed model equations as if they were regular records. Three scenarios were compared to a situation without pre-selection. Each scenario corresponded to a different type and/or amount of information included in the evaluation: actual performances of selected young sires only or combined with de-regressed genomic pseudo-performances, for the selected or all the candidate sires. Bias and accuracy of BLUP evaluations were measured for each scenario.

Populations and cohorts of the study

In their study [10], Patry and Ducrocq used actual pedigree records and records from the 2008 national type trait evaluations for the Holstein breed in France to simulate breeding values of selection candidates. The animals of interest were defined as the youngest progeny-tested bulls with no second crop daughters, hereafter called young sires (YS). Their daughters and the dams of their daughters were also known. Two populations were considered for BLUP evaluations, one in which progeny testing was carried out (CONTROL population) and one reflecting genomic pre-selection in the last generation (GPS population). To mimic genomic pre-selection among YS, GEBV were generated together with true breeding values (TBV) in the GPS population. GEBV of full-sib families of candidate sires were generated. Among each full-sib family, it was assumed that the sib with the highest GEBV was selected, while the remaining full-sibs were culled. In the CONTROL population, only TBV were simulated for YS. As with the real datasets, only selected sires had daughters, and their performances were simulated. In the current study, as in [10], the same cohorts and sets of data were used, including GEBV and TBV for all candidate sires, and performances for their daughters.

Data generation: TBV, GEBV, performances

For young sires, TBV and GEBV were simulated jointly (in the GPS population) from multivariate normal distributions and conditional on parent average (EBV before including progeny information). Variances and covariances of the distributions depended on the genetic variance of the trait and reliabilities of genetic and genomic evaluations. Direct genomic reliability and pedigree reliability were distinguished. Reliability of GEBV was defined as a combination of genomic and pedigree-based information.

Pedigree reliabilities were obtained from the true data analysis before including progeny information. Direct genomic reliabilities were computed assuming the genomic contribution contributed n additional daughter records. Various values of n were used in the simulations. Daughter performances were computed using estimated fixed effects from the true data analysis, simulated TBV of YS, and the distribution of dam EBV. For more details, see Patry and Ducrocq [10]. Simulations were replicated 50 times.

Estimation of breeding values

Breeding values were estimated for all the animals in both populations, CONTROL and GPS, based on daughter performances and pedigree-based information and using BLUP applied to a single-trait animal model. In the CONTROL population, EBV of YS were unbiased [10]. In the GPS population, only pre-selected YS had daughters and therefore, only their performances were available for the BLUP evaluation. Genomic pre-selection was not taken into account in the estimation of breeding values by BLUP and EBV of YS were shown to be biased [10].

Computation of de-regressed GEBV

To account for the genomic selection step in BLUP evaluations at the national level, GEBV were de-regressed as described in the following paragraph and weighed by the increase in reliability due to genomic information. Estimated breeding values $\hat{\mathbf{a}}$ are usually obtained as solutions of the MME:

$$\begin{bmatrix} \mathbf{X'R^{-1}X} & \mathbf{X'R^{-1}Z} \\ \mathbf{Z'R^{-1}X} & \mathbf{Z'R^{-1}Z + A^{-1}\alpha} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X'R^{-1}y} \\ \mathbf{Z'R^{-1}y} \end{bmatrix} \quad (1)$$

where $\boldsymbol{\beta}$ and $\mathbf{a}$ are vectors of fixed effects and breeding values, $\mathbf{A}$ is the additive genetic relationship matrix, $\mathbf{X}$ and $\mathbf{Z}$ are incidence matrices assigning observations to effects, and α is the variance ratio between residual and genetic variance ($\alpha = \sigma_e^2/\sigma_a^2$). From (1), EBV $\hat{\mathbf{a}}$ can be computed from:

$$\mathbf{(Z'R^{-1}Z + A^{-1}\alpha)\hat{a} = Z'R^{-1}(y - X\hat{\beta})} \quad (2)$$

This equation is obtained after correction for the breeding value of their dam and absorption of each daughter equation, such that only equations corresponding to sires and their ancestors are left.

In a regular de-regression procedure, as described by Jairath et al. [11], the **EDP** vector is obtained from the right hand side of:

$$\mathbf{(EDC + A_s^{-1}\alpha)\hat{a}_s = EDC.EDP} \quad (3)$$

where **EDC** is a diagonal matrix of Effective Daughter Contributions with element EDC_i representing the amount of information coming from daughter phenotypes for each sire i. **EDP** is a vector of de-regressed proofs also called Effective Daughter Performances; and $\mathbf{A_s}$ and $\mathbf{a_s}$ are the numerator relationship matrix and the vector of breeding values of the sires and their ancestors. Assuming that $\mathbf{a_s}$ is known from the solution of (1) or (2), we have:

$$\mathbf{EDP = (EDC)^{-}(EDC + A_s^{-1}\alpha)\hat{a}_s} \quad (4)$$

Equation (4) can be adapted to compute for each genotyped sire i, a "genomic" pseudo-performance EDP_i^g, similar to the effective daughter performance EDP_i .

Let $\Delta\mathrm{Rel}_i$ be the increase in reliability of DGV(Direct Genomic Value) or GEBV for sire i compared to its classical EBV. It will be referred to as the "direct genomic reliability": $\Delta\mathrm{Re}\,\mathrm{l}_i = \dfrac{\mathrm{EDC}_i^g}{\mathrm{EDC}_i^g + \mathrm{k}}$ or equivalently: $\mathrm{EDC}_i^g = \dfrac{\mathrm{k}\Delta\mathrm{Re}1_i}{1 - \Delta\mathrm{Re}1_i}$ where EDC_i^g is the "genomic" effective daughter contribution, $\mathrm{k} = \dfrac{4 - \mathrm{h}^2}{\mathrm{h}^2}$ and h^2 is the heritability of the trait. Replacing in $\mathbf{a_s}$ equation (4) by $\mathbf{g}$, the vector of GEBV, it follows that the vector $\mathbf{EDP^g}$ of genomic pseudo-performances is the solution of:

$$\mathbf{(EDC^g + A_s^{-1}\alpha)\,\hat{g} = EDC^g.EDP^g} \quad (5)$$

Note that vector $\mathbf{g}$ does not only include GEBV for genotyped animals but also GEBV for non-genotyped ancestors. $\mathbf{g}$ was split into two vectors ($\mathbf{g}_g$, $\mathbf{g}_{ng}$) distinguishing genotyped animals (g) from non genotyped ones (ng). After appropriate reordering of rows and columns, let:

$$\mathbf{A_s^{-1}} = \begin{bmatrix} \mathbf{A^{gg}} & \mathbf{A^{gng}} \\ \mathbf{A^{ngg}} & \mathbf{A^{ngng}} \end{bmatrix} \quad (6)$$

Assuming EDP_i^g equal to zero for non genotyped sires, vector $\mathbf{g}_{ng}$ is computed solving the following equation:

$$\mathbf{A^{ngng}\hat{g}_{ng} = -A^{ngg}\hat{g}_g} \quad (7)$$

This de-regression procedure removes the parent average effect. Therefore, either GEBV which include a residual polygenic effect or DGV can be used in $\mathbf{g}$.

To be able to include the genomic pseudo-performances in a national genetic evaluation, sire $\mathbf{EDP^g}$ and $\mathbf{EDC^g}$ must be adapted to an animal model, where the sire variance used in a sire model is replaced by the additive genetic variance. This is done by multiplying $\mathbf{EDP^g}$ by 2 and by multiplying $\mathbf{EDC^g}$ by $\dfrac{\alpha}{\mathrm{k}}$ [12].

Inclusion of genomic pseudo-performances into BLUP evaluations

For the GPS population, three different datasets of performances were created to obtain BLUP evaluations,

leading to three scenarios to account for genomic pre-selection at the national level:

1. BLUP evaluations included only one type of phenotypes for the YS, i.e. the simulated performances of their daughters. These "actual" phenotypes were available only for the YS which were pre-selected based on their genomic information. Thus culled YS were not included in the evaluation. We called this scenario "GPS_no" and has been shown to result in biased EBV [10].

2. BLUP evaluations included two types of phenotypes for the YS, i.e. the simulated performances used in scenario GPS_no and the genomic pseudo-performances $\mathbf{EDP}^{g}$, i.e. the de-regressed GEBV derived above, but $\mathbf{EDP}^{g}$ were only available for the selected candidates. This scenario was called "GPS_sel".

3. BLUP evaluations used the same two types of phenotypes available for the pre-selected YS as in the previous scenario but this time, the genomic pseudo-performances $\mathbf{EDP}^{g}$ were also included for candidates culled after the genomic pre-selection step. This scenario was called "GPS_all". Hence all candidate sires have an associated pseudo-performance.

Sets of parameters

Different levels of trait heritability, different proportions of retained young sires after genomic selection and different accuracies of GEBV were used to define several parameter sets, as in Patry and Ducrocq [10]. Thus, two type traits, udder depth (UD) and foot angle (FA), were considered because of their contrasted heritabilities (0.36 versus 0.14). The genetic variance was 0.25 for UD and 0.14 for FA. Seven hundred and ninety-nine selected YS and a total of 40,222 daughters with UD records and 601 selected YS and their 31,976 daughters with FA records were identified. Two proportions of selected YS were tested: 10% and 25%. For example, when 10% of YS were retained after genomic selection, 7,990 pairs of TBV and GEBV for UD were simulated to identify, after proper ranking, 799 selected YS and 7,191 culled YS. We assumed an initial value of 10 effective daughter records so that the direct genomic reliability was 0.50 for UD and 0.26 for FA. Because of the lower heritability of FA, we also tested a value of 26 EDC to achieve a direct genomic reliability of 0.50, as for UD. See Table 1 for the definition of all the parameter sets. Depending on the set of parameters and on the scenario (GDP_no, GDP_sel and GDP_all), a different number of actual daughter performances and genomic pseudo-performances were included, see Table 2.

Statistical analysis of the data

National BLUP evaluations were performed in the four situations presented in Table 2. Breeding values were estimated in the CONTROL population and under the three scenarios in the GPS population (GPS_no, GPS_sel, GPS_all). Before further statistical analysis of the resulting EBV, all EBV were expressed in genetic standard deviation units of the trait (σ_G). The mean Mendelian sampling term was estimated as the mean difference between the young sires' EBV and their parent average across all the YS included in each scenario. This estimate indicated how much the usual MME assumption of zero expectation for the Mendelian sampling term was violated. As in Patry and Ducrocq [10], three indicators were used to assess the quality of BLUP evaluations and were compared among the four scenarios: bias, true reliability (ρ^2) and mean square error (MSE), as defined below. Let a_i and a_i be respectively the TBV and EBV of each young sire i in each replicate r.

$$bias = \frac{1}{50}\sum_{r=1}^{50}\left(\frac{1}{n}\sum_{i=1}^{n}(\widehat{a_i} - a_i)\right) \tag{8}$$

$$\rho^2 = \left(\frac{1}{50}\sum_{r=1}^{50}\frac{\mathrm{cov}(\widehat{a_r}, a_r)}{\sqrt{\mathrm{var}(\widehat{a_r})\mathrm{var}(\widehat{a_r})}}\right)^2 \tag{9}$$

$$MSE = \frac{1}{50}\sum_{r=1}^{50}\left(Var(\widehat{a_r} - a_r) + (\widehat{a_r} - a_r)^2\right) \tag{10}$$

True reliability and MSE characterize the accuracy of BLUP evaluations. Statistics were computed for two groups of interest, the young sires and their daughters and averaged over the 50 replicates. For both groups and each scenario, they were calculated for all animals actually included in the BLUP evaluations: both eliminated and selected candidates were analysed in the CONTROL and GPS_all scenarios whereas only selected candidates were included in the analysis of GPS_no and GPS_sel scenarios.

Results

Including information on all the selection candidates avoids pre-selection bias

To illustrate the bias process and the approach to account for pre-selection, only the results for the evaluation of UD (h^2 = 0.36) when 25% of the YS were retained based on their GEBV will be presented. In the CONTROL population, the EBV of YS were unbiased since all the selection candidates were included in the BLUP evaluation (Table 3): both the mean Mendelian sampling estimate and the mean difference between true and estimated breeding values were not significantly different from zero. In contrast, the mean Mendelian sampling estimate and the bias were significantly different from zero, true reliability decreased and MSE increased, when genomic pre-selection of sires was applied (GPS population) but not accounted for in the evaluation

Table 1 Size of the cohorts according to different levels of heritability and genomic selection intensity

Proportion of selected young sires		10%			25%	
Traits	**Full-sibs family size**	**Number of selected young sires (and their daughters)**	**Number of culled young sires (without daughters)**	**Full-sibs Family size**	**Number of selected young sires (and their daughters)**	**Number of culled young sires (without daughters)**
Udder depth (h^2 = 0.36)	7990	799 (40222)	7191	3196	799 (40222)	2397
Foot Angle (h^2 = 0.14)	6010	601 (31976)	5409	2404	601 (31976)	1803

(GPS_no scenario). When genomic pseudo-performances were included for selected sires only (GPS_sel scenario), the true reliability of BLUP evaluations increased compared to the scenario GPS_no due to the explicit addition of genomic information to the traditional pedigree and performance information. The MSE also decreased, indicating that the quality of BLUP evaluations was improved. However, the bias was still significantly different from zero (Table 3). The genomic selection process was completely accounted for only when genomic pseudo-performances for culled sires were also included in the evaluation model (GPS_all scenario). In this case, the mean Mendelian sampling estimate and the bias of the cohort of selected sires were not significantly different from zero. Including de-regressed GEBV for all YS in the evaluation model as in GPS_all scenario not only accounted for genomic pre-selection, contrary to the GPS_sel scenario, but also increased accuracy of BLUP evaluations compared to the GPS_no scenario.

Influence of heritability and pre-selection intensity

Previous research showed that, when the trait heritability is lower or the genomic pre-selection intensity is higher, the relative magnitude of the bias due to genomic selection increases when the genomic pre-selection intensity is not accounted for in the evaluation model [10]. The average bias and MSE are presented in Table 4 for YS and in Table 5 for their daughters for different combinations of trait heritability and genomic pre-selection intensity levels and when selection based on genomic information is fully (GPS_all) or not accounted for (GPS_no). For the YS cohort (Table 4) in the GPS_no scenario, the bias ranged from -0.146 to -0.338 σ_G, and from -0.03 to 0 σ_G in the GPS_all scenario. In the latter case, the bias was also almost zero in the cohort of daughters (Table 5). Regardless of the magnitude of the initial bias for YS or their daughters, including genomic pseudo-performances for all the selection candidates provided the MME with sufficient information on the selection process to effectively reduce the bias.

Table 2 Number and type of performances available in BLUP evaluations for the four tested scenarios

Proportion of sires retained after genomic selection		25%		10%	
Performances		**Actual daughter records**	**Genomic pseudo-performances**	**Actual daughter records**	**Genomic pseudo-performances**
UD[d] trait	After progeny testing	40222	0	40222	0
	After genomic pre-selection: GPS_no[a]	40222	0	40222	0
	GPS_sel[b]	40222	799	40222	799
	GPS_all[c]	40222	3196	40222	7990
FA[e] trait	After progeny testing	31976	0	31976	0
	After genomic pre-selection: GPS_no[a]	31976	0	31976	0
	GPS_sel[b]	31976	601	31976	601
	GPS_all[c]	31976	2404	31976	6010

[a]genomic pre-selection of young sires but no inclusion of genomic pseudo-performances; [b]genomic pre-selection of young sires and genomic pseudo-performances were included for selected young sires; [c]genomic pre-selection of young sires and genomic pseudo-performances were included for all candidate sires; [d]udder depth; [e]foot angle

Table 3 Quality of BLUP evaluations of young sires for udder depth after a 25% genomic pre-selection

Scenarios	Mendelian sampling estimate (in σ_g^d)	Bias (in σ_g^d)	True reliability	Mean square error
CONTROL	-0.001 (ns)	0.002 (ns)	0.756	0.183
GPS_no[a]	0.304 (***)	-0.146 (***)	0.727	0.188
GPS_sel[b]	0.188 (***)	-0.138 (***)	0.763	0.165
GPS_all[c]	-0.003 (ns)	-0.019 (ns)	0.760	0.150

H0 = {μ = 0}: ns = non significant (p > 0.001); *** = p-value < 0.001; [a]genomic pre-selection of young sires but no inclusion of genomic pseudo-performances; [b]genomic pre-selection of young sires and genomic pseudo-performances were included for selected young sires; [c]genomic pre-selection of young sires and genomic pseudo-performances were included for all candidate sires; [d]genetic standard deviation of the trait

Impact of genomic evaluation accuracy

In the previous situations, we considered diagonal values of $\mathbf{EDC}^g$ of 10 for UD and 26 for FA. In Table 6, we compared these results for FA with a situation where diagonal values of $\mathbf{EDC}^g$ were assumed to be 10 instead of 26; hence the accuracy of the genomic evaluations was assumed to be lower. In this case, the expected genetic gain genetic trend was smaller and selection was less efficient. As a result, the bias due to not accounting for pre-selection (GPS_no) was smaller than with an $\mathbf{EDC}^g$ of 26. However, the bias was also less reduced by including genomic pseudo-performances for selected YS (GPS_sel) when EDCg was equal to 10. This illustrates the fact that the accuracy of GEBV is a key element when including genomic performances for all candidates in the evaluation model to account for bias due to genomic pre-selection.

Discussion

The inclusion of a genomic pseudo-performance, i.e. a de-regressed GEBV, for all genotyped candidates reduced the GEBV bias to (almost) zero in most simulated situations, regardless of the genomic selection intensity. Inclusion of genomic pseudo-performance resulted in a better description of the genetic characteristics of the population of candidates. Consequently, the overall average Mendelian sampling term had a zero expectation and the classical assumptions of the BLUP model were more closely met. However, the results showed that the effectiveness of this approach depended on the quality of genomic evaluations. This approach was more effective for traits with a higher heritability or for genomic evaluations with a higher accuracy. As expected, adding genomic data increased the amount of information contributed to the genetic evaluation and this information was distributed to relatives through the additive relationship matrix. In fact, including genomic pseudo-performance is not as straightforward as adding regular performance to BLUP evaluations [13]: obviously, accuracy of EBV increases as the number of daughters increases but this is not always the case with an increasing number of genotyped animals. Indeed, genotyped parents correctly add information to non-genotyped progeny and genotyped progeny contribute information to non-genotyped parents but the total amount of additional information from genotyped relatives cannot exceed the gain in accuracy from genotyping the animals themselves [8]. Furthermore, if a progeny and its sire are both genotyped, the progeny genotype does not provide any additional information to the sire and vice versa [6]. Thus including without care genomic pseudo-performances for both the sire and its progeny will result in double counting genomic contributions, once directly, and once via relatives through the additive relationship matrix [8]. Therefore, BLUP evaluations must account for such data redundancy.

In this study, only YS were genotyped and we implicitly assumed that none of their sires were from the reference population, hence avoiding the issue of redundant genomic information and overestimated reliability of genomic evaluation [14]. However, in a more realistic case, the weight of genomic information might be overestimated by $\mathbf{EDC}^g$ and a tailored reduction of $\mathbf{EDC}^g$ should be implemented. Nevertheless, despite the simplified assumptions and computations, the approach used was

Table 4 Quality of BLUP evaluations with or without accounting for pre-selection in the cohort of selected young sires

Heritability	Proportion of selected young sires	Bias (in σ_g^c)		Mean squared error	
		GPS_no[a]	GPS_all[b]	GPS_no[a]	GPS_all[b]
0.36 (UD[d] trait)	10%	-0.227 (***)	-0.030 (*)	0.217	0.157
	25%	-0.146 (***)	-0.019 (ns)	0.188	0.150
0.14 (FA[e] trait)	10%	-0.338 (***)	-0.020 (ns)	0.364	0.222
	25%	-0.214 (***)	-0.011 (ns)	0.305	0.229

H0 = {μ = 0}: ns = non significant (p > 0.001); *** = p-value < 0.001; [a]genomic pre-selection of young sires but no inclusion of genomic pseudo-performances; [b]genomic pre-selection of young sires and genomic pseudo-performances were included for all candidate sires; [c]genetic standard deviation of the trait; [d]udder depth; [e]foot angle

Table 5 Quality of BLUP evaluations with or without accounting for pre-selection in the cohort of daughters of the selected young sires

Heritability	Proportion of selected young sires	Bias (in σ_g[c])		Mean squared error	
		GPS_no[a]	GPS_all[b]	GPS_no[a]	GPS_all[bb]
0.36 (UD[d] trait)	10%	-0.074 (***)	-0.009 (ns)	0.547	0.541
	25%	-0.044 (**)	-0.002 (ns)	0.544	0.540
0.14 (FA[e] trait)	10%	-0.144 (***)	-0.010 (ns)	0.685	0.662
	25%	-0.092 (***)	-0.006 (ns)	0.674	0.660

H0 = {μ = 0}: ns = non significant (p > 0.001); *** = p-value < 0.001; [a]genomic pre-selection of young sires but no inclusion of genomic pseudo-performances; [b]genomic pre-selection of young sires and genomic pseudo-performances were included for all candidate sires; [c]genetic standard deviation of the trait; [d]udder depth; [e]foot angle

shown to be promising and demonstrated that including information on culled candidates is essential.

With the addition of genomic information, inflated reliabilities have been reported regardless of the method used to blend genomic and traditional information: the selection index approach [2,14], the single-step approach [4], or the current approach, which was initially proposed by Ducrocq and Liu [6]. Some strategies have been suggested to prevent the reliability of genotyped animals from approaching 1. Ducrocq and Liu [6] have proposed an iterative approach adapted from the information source method [15] to compute reliability from genomic information. In their situation, the **EDC**[g] were derived under constraints such that the final genomic contribution to reliabilities was bounded. The reliabilities of GEBV appeared to be reasonable. However, the issue was not completely solved since reliabilities were still overestimated for sires with many genotyped progeny [6]. Mäntysaari and Stràndén have proposed to use a multi-trait evaluation to combine DYD and DGV, where DGV are treated as an indicator trait with a high correlation to the considered trait. Then, reliabilities of GEBV are naturally bounded to the square of this correlation so that genomic relationships are less overestimated. Such a correlation between EBV and DGV or GEBV could be estimated following the method proposed by Kachman [17] and implemented by MacNeil et al.[18].

The single-step approach [4,5] offers an appealing solution in the sense that genomic, phenotypic and pedigree information are analyzed simultaneously. However, unless it is assumed that all the genetic variation is described by the SNP markers, these procedures face the problem of finding an appropriate weighting of genomic and pedigree-based information [4,5]. In some studies, the lack of independency between the three sources of information (genomic, phenotypic, pedigree based) has been considered through a scaling of the residual variance [16,19] but only approximate solutions have been developed so far. Further appropriate developments are necessary to better compute **EDC**[g] and to improve the method of including genomic performances in BLUP evaluation to account for bias due to genomic pre-selection. The approach presented here involves an additional step, before running national BLUP evaluations, i.e. computation of genomic pseudo-performances. This step is easy to implement as de-regression is commonly used, like in routine international genetic evaluations [11]. This method has several key advantages. First, it is independent from the methodology used to predict genomic EBV (GBLUP, Bayesian methods, etc), secondly, it can be applied to different evaluation models without further developments and, finally, the size of the genotyped population is not a constraint.

With the current breeding schemes in dairy cattle, a period of about four years is necessary between the genomic pre-selection step and the introduction of the first records of daughters in BLUP evaluations. Since genomic selection has begun more than two years ago in several

Table 6 Effect of accuracy of genomic evaluations on BLUP evaluations for foot angle in the cohort of selected young sires

EDC[g] [d]	Proportion of selected young sires	Bias (in σ_g[c])		Mean squared error	
		GPS_no[a]	GPS_all[b]	GPS_no[a]	GPS_all[b]
10	10%	-0.249 (***)	-0.098 (***)	0.339	0.270
	25%	-0.155 (***)	-0.054 (**)	0.299	0.257
26	10%	-0.338 (***)	-0.020 (ns)	0.305	0.222
	25%	-0.214 (***)	-0.011 (ns)	0.364	0.229

H0 = {μ = 0}: ns = non significant (p > 0.001); *** = p-value < 0.001; [a]genomic pre-selection of young sires but no inclusion of genomic pseudo-performances; [b]genomic pre-selection of young sires and genomic pseudo-performances were included for all candidate sires; [c]genetic standard deviation of the trait; [d]effective daughter contribution from genomic EBV

countries, the first biased evaluations may occur within the two next years. Thus the need to implement an easy to apply approach to account for genomic pre-selection is urgent. The approach proposed here requires only limited modifications (if any) of the existing national evaluation software. However, further work is needed to control the dependency between BLUP evaluations and genomic evaluations. To account for genomic pre-selection, EBV must include genomic information and these unbiased EBV are then used as input for future equations for genomic predictions. The issue is that genomic information will be double counted when computing GEBV. One way to circumvent this problem would be to iterate between the classical genetic and genomic evaluations.

Two alternatives, both potentially problematic, are possible: on the one hand, genomic pre-selection of young sires leads to biased EBV and therefore to biased DYD which are then used to update genomic predictions. On the other hand, incorporating genomic records into national BLUP evaluations inflates the accuracy of BLUP EBV of some animals and makes classical genetic and genomic evaluations dependent from each other. Thus, a compromise has to be found between the use of biased EBV on one side, and double counting of genomic information and overestimation of reliabilities on the other side.

In this study, the underlying context was rather optimistic. In particular, it was assumed that all data from selected and culled candidates were available at the national level. For example, the use of pre-selected bulls from foreign breeding schemes was not considered. Moreover, in the context of national and international competition, breeding companies may be reluctant to release information on their selection strategy and objectives, and may not be willing to share data on culled animals. Our study clearly shows that this would be very detrimental for at least three reasons: first, EBV of pre-selected bulls would be underestimated; secondly, the resulting bias would be transferred to the rest of the population (e.g., daughters) in an uncontrolled way; and finally, genomic predictions using results from these biased evaluations would be sub-optimal. Therefore, it is essential that information originating from current implementations of genomic selection (GEBV of all animals, or at least selection differentials) at least be shared at the national level. Ignoring genomic pre-selection at the national level impacts national EBV and, as a consequence, international EBV too. We are currently investigating to what extent the transmission of biased or unbiased national EBV for selected bulls only could bias international genetic evaluations.

Conclusions

There is an urgent need to account for genomic pre-selection of young sires before their national EBV become biased. Based on a real dairy cattle dataset, breeding values were generated in the last generation of sires to mimic genomic pre-selection. In this study, including a genomic pseudo-performance based on GEBV for all the selection candidates strongly reduced or removed biases, regardless of their magnitude. However, this approach does not account for some potential overestimation of the weight that is placed on genomic information and for dependency of genetic and genomic evaluations. Thus, the proposed method may need further improvement, but in the short term, it makes possible to implement a simple and general procedure that accounts for these new selection practices in BLUP evaluations at the national level. In addition, this approach provides an alternative method to combine genomic, phenotypic and pedigree data in multiple steps procedures which is easy to understand and implement.

Acknowledgements

Financing of the AMASGEN project (Jouy-en-Josas, France) by Agence Nationale de la Recherche and APISGENE is gratefully acknowledged. We would like to thank the reviewers for their comments and corrections.

Author details

[1]INRA, UMR 1313 Génétique Animale et Biologie Intégrative, F-78350 Jouy-en-Josas, France. [2]UNCEIA, Département de Génétique, 149, rue de Bercy, F-75595 Paris Cedex, France.

Authors' contributions

CP implemented the proposed methodology and participated in the results analysis.
VD conceived the study, participated to its implementation and in the analysis. CP and VD were both involved in drafting the manuscript. They both read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 18 March 2011 Accepted: 18 August 2011
Published: 18 August 2011

References

1. Goddard M, Hayes B: **Genomic selection.** *J Anim Breed Genet* 2007, **124**:323-330.
2. VanRaden P: **Efficient methods to compute genomic predictions.** *J Dairy Sci* 2008, **91**:4414-4423.
3. Misztal I, Legarra A, Aguilar I: **Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information.** *J Dairy Sci* 2009, **92**:4648-4655.
4. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ: **Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score.** *J Dairy Sci* 2010, **93**:743-752.
5. Christensen O, Lund M: **Genomic prediction when some animals are not genotyped.** *Genet Sel Evol* 2010, **42**:2.
6. Ducrocq V, Liu Z: **Combining genomic and classical information in national BLUP evaluations.** *In Proceedings of the 2009 Interbull meeting: 21-24 August 2009; Barcelona.* , Edited by Interbull bulletin 2009 2009, **40()**:172-177.
7. Vitezica ZG, Aguilar I, Misztal I, Legarra A: **Bias in genomic predictions for populations under selection.** *Genet Res (Camb)* 2011, 1-10.
8. VanRaden P, Sullivan P: **International genomic evaluation methods for dairy cattle.** *Genet Sel Evol* 2010, **42**:7.
9. Henderson CR: **Best Linear Unbiased Estimation and Prediction under a selection model.** *Biometrics* 1975, **31**:25.

10. Patry C, Ducrocq V: **Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle.** *J Dairy Sci* 2011, **94**:1011-1020.
11. Jairath L, Dekkers JC, Schaeffer LR, Liu Z, Burnside EB, Kolstad B: **Genetic evaluation for herd life in Canada.** *J Dairy Sci* 1998, **81**:550-562.
12. Fikse W, Banos G: **Weighting factors of sire daughter information in international genetic evaluations.** *J Dairy Sci* 2001, **84**:1759-1767.
13. Harris BL, Johnson DL, Spelman RJ: **Genomic selection in New Zealand and the implications for national genetic evaluation.** In *Proceedings of the 36th International Committee for Animal Recording Biennial Session: 16-20 June 2008; Niagara Falls. Volume 13.* ICAR Technical series; 2008:325-330.
14. Van Raden P, Van Tassell C, Wiggans G, Sonstegard T, Schnabel R, Taylor J, Schenkel F: **Invited review: reliability of genomic predictions for North American Holstein bulls.** *J Dairy Sci* 2009, **92**:16-24.
15. Harris B, Johnson D: **Approximate Reliability of Genetic Evaluations Under an Animal Model.** *J Dairy Sci* 1998, **81**:2723-2728.
16. Mäntysaari EA, Strandén I: **Use of bivariate EBV-DGV model to combine genomic and conventional breeding value estimations.** *Proceedings of the 9th World Congress of Genetics Applied to Livestock Production: 1-6 August 2010; Leipzig* 2010.
17. Kachman S: **Incorporation of marker scores into national genetic evaluations.** *9th genetic prediction workshop; Kansas City, Missouri* federation Bi; 2008, 92-98.
18. MacNeil MD, Nkrumah JD, Woodward BW, Northcutt SL: **Genetic evaluation of Angus cattle for carcass marbling using ultrasound and genomic indicators.** *Journal of Animal Science* 2010, **88**:517-522.
19. Edel C, Emmerling R, Götz K-U: **A modification of Van Raden's index for the blending of genomic breeding values.** *Proceedings of the 9th World Congress of Genetics Applied to Livestock Production: 1-6 August 2010; Leipzig* 2010.

doi:10.1186/1297-9686-43-30
Cite this article as: Patry and Ducrocq: **Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle.** *Genetics Selection Evolution* 2011 **43**:30.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central

Erratum – article II

1) Equation (9) should be modified: $\rho^2 = \frac{1}{r}\sum_{r=1}^{50}\left(\frac{\mathrm{cov}(\hat{a}_r, a_r)}{\sqrt{\mathrm{var}(\hat{a}_r)\,\mathrm{var}(a_r)}}\right)^2$

2) Table 6 should be replaced by the following:

Effect of precision of genomic evaluations on BLUP evaluations for foot angle in the cohort of selected young sires

EDC_g	Proportion of selected young sires	Bias (in σ_g)		Mean Squared Error	
		GPS_no	GPS_all	GPS_no	GPS_all
10	10%	-0.249 (***)	-0.098 (***)	0.339	0.270
	25%	-0.155 (***)	-0.054 (**)	0.299	0.257
26	10%	-0.338 (***)	-0.020 (ns)	0.364	0.222
	25%	-0.214 (***)	-0.011 (ns)	0.305	0.229

CHAPTER 6 - Bias propagation in international evaluations due to genomic selection

National and international genetic evaluations are interdependent:

- international genetic evaluations require classical evaluations provided by countries to be de-regressed and used as phenotypes in MACE analyses;
- national classical evaluations need to include MACE breeding values to consider foreign sires in national selection schemes;
- national genomic evaluations also require MACE breeding values as phenotypes to estimate markers effects from a large multinational reference population (RP).

First, Interbull is a central player in the genetic evaluation process and it is important to describe how this organization is affected by the wide implementation of genomic selection.

Then, the reasons for bias in international evaluations due to genomic selection will be presented.

Finally, Patry *et al.* (submitted) have simulated the effects of genomic selection on the international evaluations to assess bias and measure the impacts at the international level.

6.1. The roles of Interbull: changes due to genomic selection

6-1.1. THE ROLES OF INTERBULL

The Interbull Centre was established in 1991 under contract with the Swedish University of Agricultural Sciences in Uppsala, Sweden. Interbull has three historical roles:

- International estimated breeding values on various country scales

One role is to provide a tool for international comparison of dairy sires. Using the MACE method described in chapter 2, each sire is evaluated on all country scales. Today, 28 countries are sending pedigree and national breeding value estimates to Interbull. International EBV are delivered for 6 breeds (Brown Swiss, Guernsey, Holstein, Jersey, Red Dairy Cattle, Simmental) and 7 groups of traits (production, conformation, udder health, longevity, female fertility, calving and workability). To be consistent with national genetic systems, 3 MACE routine runs are carried out each year (April, August, December) and 2 test runs (January and September) are realized to account for changes in national genetic systems. MACE EBV are a tool to get early and accurate EBV for foreign bulls but also to establish an international ranking. International genetic evaluations have helped international trade of semen.

- Validation of national genetic evaluations

Another role of Interbull is to validate national evaluations. Indeed, it provides a quality standard assurance for national evaluations before countries are allowed to participate to international evaluations. Boichard *et al.* (1995) proposed three different methods to test trends in genetic gain (tests I, II, and III). These assume no changes in EBV after inclusion of information from additional lactations and/or daughters in genetic evaluations. More recently, Fikse *et al.* (2003) proposed a method to test trend in genetic variances, relying on the estimation of the MS distribution (Test IV). Test IV is still under developments (Lidauer *et al.*, 2007, Tyriserva *et al.*, 2011) and is not yet routinely implemented.

- An European reference

The Interbull Centre is also the European Union reference laboratory for animal production, i.e., the body responsible for stimulating uniform testing methods and for the assessment of the characteristics of pure bred animals in cattle.

6-1.2. NEW INTERBULL'S ROLES IN THE GENOMIC ERA

With the implementation of genomic selection, some of the missions of Interbull have been under review.

- Validation of national genomic evaluations

Interbull continues to validate national evaluations including the classical evaluations but also the genomic ones. Mäntysaari *et al.* (2010) adapted the previous methods (test III) to genomic breeding values, leading to test the consistency of the genetic trend captured by GEBV and the variation of GEBV. The statistical tests are still under discussion but the validation of genomic validation is critical and is already part of the new missions of Interbull.

- International estimated breeding values including genomic information

A new aim of Interbull is to deliver international genomic breeding values: Interbull is trying to adapt the MACE method to analyze together national GEBV. Classical MACE assumes independent residuals since each cow is recorded in only one country. On the contrary, one bull can be genotyped and evaluated in several countries at the same time so that it generates non null residual co-variances corresponding to redundant genomic information. GMACE and simplified GMACE (S-GMACE) have been proposed (Van Raden and Sullivan, 2010) to combine classical EBV with GEBV in a same international analysis. This has been tested by Interbull since 2010. However, results are not yet satisfying for countries implementing genomic evaluations: the methods do not take full advantage of the increase of information brought by the genotypes. GMACE is planned to be a short-term alternative before finding a consensus on international genomic breeding values. Nevertheless, the shorter productive life of sires and the lower availability of their semen make exports more difficult and importance of an international ranking based on GMACE solutions might be less critical.

Three groups of countries participating to Interbull with different needs can now be distinguished (Interbull, 2011):

- Countries with their own bull selection schemes and genomic evaluation system;
- Semen importing countries with their own genomic evaluation;
- Semen importing countries without a genomic evaluation system.

To account for their differentiated needs, Interbull has now to maintain and adapt its usual missions but the importance of MACE is no longer the same for all countries.

6.2. The importance of international evaluations in the genomic era

International evaluations are the only way to properly rank bulls from various countries on a same country scale. This has been beneficial for all countries with genetic evaluations, i.e., for importing as well as for exporting countries.

In January 2009, Interbull organized a workshop dedicated to the integration of genomic information in the (inter)national genetic evaluations. At that time, the necessity to deliver classical estimated breeding values on the different country scales was questioned. With the implementation of genomic evaluations, sires could also be ranked according to their GMACE values.

It was especially recognized that MACE breeding values are still crucial to create (or increase the size of) multinational RP. In fact, it is essential to share not only genotypes but also phenotypes at an international level to achieve higher levels of accuracy of genomic evaluations (Hayes *et al.*, 2009). Two main consortia have been created. On the one hand, Eurogenomics gathers France, Germany, The Netherlands, the Nordic countries, and since September 2011, Spain. On the other hand, the North-American consortium brings together data from Canada, the USA, and more recently from the UK and Italy. For this purpose, EBV expressed on the same scale, i.e., MACE EBV, are used (after de-regression) as phenotypes in the estimation of SNP effects. This places again Interbull as a key player in the genetic evaluation process but for countries with own genomic evaluation only.

With the genomic era, the international organization of the dairy cattle breeding has become unbalanced: some countries are implementing genomic evaluations and selection and others not. Hence, it has become more difficult for Interbull to deliver fair international comparisons. The role of Interbull has changed but new players also emerged: consortia have been created involving new types of collaborations based not only on data but also on methods, on software, and on sharing know-how among members. The importance of MACE breeding values has been renewed and it is essential for accurate genomic predictions to get unbiased MACE breeding values.

6.3. Risks of biases in international evaluations

6-3.1. VARIOUS STRATEGIES BETWEEN COUNTRIES TO IMPLEMENT GENOMIC SELECTION

Among the 28 countries participating to Interbull, various strategies regarding genomic selection may exist:

- Countries do not have genomic evaluations yet and selection decisions are not based on genomic information;
- Countries perform genomic selection and they account for genomic selection in classical genetic evaluations;
- Countries perform genomic selection but they do not account for genomic selection.

6-3.2. TWO MAIN SOURCES OF BIAS IN INTERNATIONAL EVALUATIONS

The first group does not generate any problem for international evaluations. This is the routine case which is supposed to be unbiased. For the second and third group, the situations are different:

- **A first issue: national proofs may be incomplete**

In the second case, estimated breeding values are unbiased. The problem is whether breeding values from selected sires are only provided to Interbull. It follows that international evaluations would be performed based on an incomplete data set. This is the same problem as the one raised at the national level after genomic selection: the selection process is not fully described and assumptions for optimal properties of the mixed linear models are violated.

- **A second issue: national proofs may not only be incomplete but also incorrect**

In the last group above, EBV are delivered for the selected candidates only, the data set is incomplete. Since EBV are biased, the risk is to propagate this bias to other relatives and foreign populations through the international genetic relationship. Moreover, any bias from incomplete or incorrect data might be expressed on the different country scales through the genetic variance-covariance matrix.

6.4. Bias assessment in international evaluations

6-4.1. GENERAL STRATEGY

The two issues of incomplete and possibly incorrect data sets were examined in a third study. By simulation, the robustness of MACE to selected data was tested as well as the propagation of bias on different scales. Bias was measured in MACE solutions of young sires going

through genomic selection. Methods and results were presented in a third article: **Patry, C., H. Jorjani and V. Ducrocq. Implementation of genomic selection at national level: impact of pre-selected and biased national BLUP evaluations on international genetic evaluations, submitted to Journal of Dairy Science on September, the 27th, 2011.**

The general strategy is the same as implemented while assessing bias at the national level after a genomic selection step. Some differences are outlined here:

- ✓ Simulations were based on real data in Holstein breed but phenotypes were for a production trait, protein yield, and data was provided by France, Germany and the USA.
- ✓ Before the implementation of genomic selection worldwide, the national proofs provided to Interbull were supposed to be complete and unbiased. The set of data used to perform international evaluations in August 2010 was considered as the CONTROL case. However, GEBV were not available and for convenience, national EBV were used as proxy of GEBV to mimic genomic.
- ✓ Selection occurred within half-sibs family: based on Mendelian sampling estimates, culled candidates were identified. MS estimates were computed as the difference between parent average and individual breeding values from international evaluations. National EBV of the culled candidates were deleted to provide Interbull with an incomplete data set.
- ✓ To mimic the effect of incorrect national proofs, bias assessments from the study at the national level were used and added as a normally distributed variable to the initial (unbiased) national proofs. This generates a set of biased national proofs for international evaluations.
- ✓ Multi-trait MACE analyses were performed assuming a sire model.
- ✓ Bias was measured as the difference between MACE solutions from the unaltered data set, i.e., after progeny testing, and MACE solutions from the altered data set, i.e., after deletion of the EBV from the culled candidates and possibly use of biased proofs for the selected ones.
- ✓ Different sources of variations were also considered: international evaluations were performed for the three populations on three different scales for two types of genomic selection strategy; simulations were replicated 10 times.

6-4.2. SUMMARY OF THE SIMULATION RESULTS

Five main conclusions can be drawn from this simulation study:

- Simulations showed that bulls from countries implementing genomic selection and providing incomplete data were clearly penalized on their own scale but also on foreign scales.
- They were even more penalized, especially the young sires after genomic selection, when these countries also provided incorrect data, i.e, biased BLUP solutions.

- Bias in national proofs of the genomically selected young sires was transmitted to other foreign young sires through the international evaluation.
- It appeared difficult to predict bias magnitude when genomic selection was implemented simultaneously in several countries according to various strategies (incomplete or incomplete and incorrect data sent to Interbull).
- Disorder among rankings and top lists is certain and may lead to changes in market shares.

Genomic selection leads to two types of concerns at the international level – incomplete and incorrect national EBV - which are both prejudicial for international evaluations, it is still better though to provide Interbull with (possibly complete) correct national proofs rather than incomplete and incorrect data.

Article III

Patry, C., H. Jorjani and V. Ducrocq.

Implementation of genomic selection at national level: impact of pre-selected and biased national BLUP evaluations on international genetic evaluations

Submitted to Journal of Dairy Science on September, the 27th, 2011.

Implementation of genomic selection at national level: impact of pre-selected and biased national BLUP evaluations on international genetic evaluations, Patry. Genomic pre-selection of young bulls is now widely implemented in dairy breeding schemes. If this pre-selection step is not accounted for in the genetic evaluations models, breeding values are estimated with bias at national level. The latter are then submitted for international genetic evaluations. Bias due to genomic pre-selection was measured for three countries after international evaluation. Young bulls from the country submitting incomplete and possibly incorrect data were highly penalized. But young bulls from all countries were affected. Missing and biased data lead to incorrect breeding values and non optimal rankings, likely to impact selection decisions and market shares.

GENOMIC SELECTION: BIASED INTERNATIONAL EVALUATIONS

Implementation of genomic selection at national level: impact of pre-selected and biased national BLUP evaluations on international genetic evaluations

Clotilde Patry[1,2§], Hossein Jorjani[3], Vincent Ducrocq[1]

[1]UMR 1313 Génétique Animale et Biologie Intégrative, Institut National de la Recherche Agronomique, 78 352 Jouy-en-Josas, France
[2]Union nationale des Coopératives d'Elevage et d'Insémination Animale, 149 , rue de Bercy, 75 595 Paris Cédex 12, France
[3] Interbull Centre, Department of Animal Breeding and Genetics SLU, Box 7023, S-75007 Uppsala, Sweden
Clotilde Patry, UMR 1313 Génétique Animale et Biologie Intégrative, Institut National de la Recherche Agronomique, 78 350 Jouy-en-Josas, France
+33 (0)134 652965, +33 (0)134 652 210, clotilde.patry@jouy.inra.fr

ABSTRACT

Genomic pre-selection of young bulls is now widely implemented in dairy breeding schemes, especially in Holstein breed. However, if not accounted for in genetic evaluation models, breeding values of genomically selected sires and their progeny can be biased. It follows that countries participating in international genetic evaluations will provide a selected and possibly biased set of national proofs to Interbull. The objective of the study was to assess bias due to genomic selection at international level. Including selected and biased national proofs in international evaluations was simulated using actual national proofs as proxy of genomically enhanced breeding values from 3 countries with a large Holstein sire population. Simulated MACE results were compared to MACE results using all available national data assumed to be unbiased. Bias was measured among young bulls genomically selected. Results were analysed by country of origin of the bulls and country scale of international genetic evaluations. Bias due to pre-selection or due to national biased proofs highly penalized young bulls from the country responsible for such biases. But it also had an impact on foreign young bulls. However, it was more difficult to predict impact on young bulls when different sources of bias were combined at the same time. All effects interact with each other. But a large amount of re-rankings is certain. This study underlines the importance of accounting for genomic selection at national level and of submitting all available data to maintain the quality of the international genetic evaluations after implementation of a genomic selection step.

Key words: Interbull, international genetic evaluation, genomic selection, selection bias

INTRODUCTION

Breeding strategies in dairy cattle are being transformed by the emergence of genetic evaluation tools combining new molecular technologies and advanced statistical analyses. Genomic selection is actually developing fast and is propagating worldwide. In 2009, only few countries computed Genomic Enhanced Breeding Values (GEBV) in Holstein breed. In 2010, data from 13 populations from 5 dairy breeds and 10 countries were provided for validation of national genomic evaluations at the International Bull Evaluation Service, Interbull (Nilforooshan et al., 2010).

Since 1994, Interbull has routinely delivered international genetic evaluations on the different country scales to promote international genetic exchanges. They are obtained using an international pedigree file and an increasing amount of data per bull having progeny in several countries. International genetic evaluations are currently computed from a sire multiple-trait across country evaluation (**MACE**) model as proposed by Schaeffer (1994). Each trait is considered to be different (but correlated) across countries with a different heritability. MACE is a mixed linear method developed to obtain for each sire as many national genetic evaluations as the number of participating countries.

The changes in national evaluation systems and the new breeding strategies due to genomic advances have now to be considered in international genetic evaluations. Thus, Interbull has to face two major challenges to maintain international comparison of dairy sires across a large number of countries. First, there is a need to develop a new evaluation methodology using genomic information to provide international breeding values of sires as early as possible. Second, the international genetic evaluations as they are currently delivered i.e. based on pedigree and phenotypes, also need to be maintained. Indeed, 18 out of the 28 countries participating in international genetic evaluations do not compute genomic evaluations yet. Furthermore, international breeding values free from genomic information are required to increase international reference population and to compute genomic prediction equations.

However, the implementation of genomic selection at national level may threaten the correctness of international genetic evaluations. They may be computed based on an incomplete data set where national proofs for culled candidates are missing or ignored. Further, selection based on direct genomic breeding values (**DGV**) or GEBV may affect the Mendelian sampling (**MS**) term. As MACE methodology relies on mixed model equations which assume that the Mendelian sampling term averages to zero, it is feared that MACE results computed from a selected sub-population could be biased (Henderson, 1975).

Patry and Ducrocq (2011b) showed that national BLUP evaluations were biased due to genomic selection of young bulls at national level. Breeding values of selected candidates were found to be systematically underestimated and such a bias may become unpredictable across generations. It was thus recommended to account for genomic selection in the computation of national BLUP breeding values to prevent them from being biased. Three approaches were proposed based on BLUP evaluations including genotyped and non-genotyped animals, either by modifying the relationship matrix (Aguilar et al., 2010, Christensen and Lund, 2010, Legarra et al., 2009, Misztal et al., 2009) or by including GEBV as de-regressed performances (Ducrocq and Liu, 2009, Patry and Ducrocq, 2011a) or as a correlated trait (Mäntysaari and Strandén, 2010) for all genotyped candidates.

For international genetic evaluations, two sources of problematic situation may now be considered. Among participating countries, some have implemented genomic selection at national level. Either they account for genomic selection at national level and deliver unbiased national breeding values to Interbull or they ignore pre-selection and deliver biased breeding values. In the first case, countries submit unbiased national proofs to Interbull but from selected sires only. In the second case, two sources of bias simultaneously hamper the correctness of international genetic evaluations: not only the data set of national proofs is incomplete but it is also biased for the sires selected based on molecular information.

Genetic evaluations based on multiple-trait model are known to be more robust than single trait evaluations (Pollak et al., 1984). Given the quick implementation of genomic selection around the world, the sensitivity of MACE methodology to selected sub-populations and biased national breeding values due to genomic selection need to be assessed. This study aims at measuring the magnitude of such a bias and its consequences on selection choices and at better understanding the mechanisms that could lead to biased MACE breeding values.

MATERIALS AND METHODS

Material

National genetic evaluation results submitted by countries to participate in the August 2010 international genetic evaluation were used to simulate genomic selection and assess its impacts at international level. In the present study, we focused on one production trait, protein yield, in the Holstein breed. The population of sires to evaluate was also limited to three large countries, denoted hereafter A, B and C. A total of 57,688 sires were considered out of about the 116,000 sires when evaluations included data from 27 countries. A, B and C delivered national genetic evaluations from single trait BLUP animal model. Genetic parameters are displayed in Table 1. Heritability values were sent by the national evaluation centres and genetic correlations among countries were computed from previous international evaluations. However, sire standard deviation and international breeding values were re-estimated using the MACE methodology according to the policies implemented at Interbull centre for routine evaluations (Interbull, 2008). For each individual, animal breeding values and parental average breeding values were delivered from international evaluations. All estimated breeding values in a given country were expressed in genetic standard deviation of that country.

General Strategy

All simulated scenarios are summarized in Table 2. Simulations were based only on national estimated breeding values (**N-EBV**, i.e. national BLUP evaluations) available at Interbull and do not include any genomic information: for young bulls (the youngest cohort of bulls with N-EBV available), it was considered that these N-EBV were equivalent to genomic breeding values (i.e. GEBV for genomically enhanced breeding values or DGV for direct genomic values) that may be or may be not sent to Interbull for some or all of them. In other words, national genomic pre-selection was mimicked by assuming that some of the N-EBV of the youngest cohort of sires were not sent to Interbull.

The effect on international evaluations of including selected data was considered first. It was assumed that pre-selection occurred in one country at a time (defining scenarios **SEL-A**, **SEL-B** and **SEL-C**) or in all countries at the same time (scenario **SEL-all**) but that this pre-selection was properly accounted for in the national evaluation, leading to unbiased national (G)EBV. These SEL scenarios represent the situation where a genetic evaluation centre did account for pre-selection but was not willing to transmit any information on bulls culled on genomic information to Interbull. The second set of scenarios (**BNP-A**, **BNP-B**, **BNP-C** and **BNP-all**), where BNP stands for Biased National Proofs, is similar to the SEL scenarios, but assumes that pre-selection was not accounted for at national level, producing biased national proofs as illustrated in Patry and Ducrocq (2011a). In the BNP-all scenario, the 3 countries implemented genomic selection and none of them did account for genomic selection at national level. Finally, an international evaluation was also considered using all the available information ("**CONTROL**" evaluations). Its results will define the reference situation. All international breeding values delivered under the SEL and BNP scenarios were compared to the CONTROL ones.

Genomic selection was assumed to be implemented among the cohort of the youngest bulls. The latter, hereafter called **YB**, were defined as the sires born between 2003 and 2006 and having only daughters in their country of origin. The bias was measured on this cohort of 7,118 YB. Among them, 2,234 had proofs in A, 1,282 in B and 3,602 in C.

Simulating the SEL Scenarios

Compared with the CONTROL case, SEL evaluations were computed based on a reduced list of sires. National proofs for culled candidates were actually missing. Assume that YB were selected based on genomic information. As mentioned before, candidates were deleted based on N-EBV used as proxys of GEBV. Selection was performed within half-sib family of YB in each country. Selection was implemented only in large families, i.e. sire families including more than 10 half-sibs. To account for the effect of reducing the quantity of information involved in international evaluations, two scenarios were looked at: within family selection was performed either according to the YB Mendelian sampling estimates (**MS SEL** scenarios) or according to a random value (**RD SEL** scenarios). MS estimates were computed as the difference between MACE parental average and MACE animal breeding values obtained in the CONTROL evaluations. Ten percent of the candidates with the highest values for the chosen criterion (MS or RD) per country were retained. Finally, 50,779 sires were evaluated in MS (or RD) SEL-A scenarios, 50,686 in MS (or RD) SEL-B scenarios, 50,917 in MS (or RD) SEL-C scenarios and 51,272 in MS (or RD) SEL-all scenarios instead of 57,688 in the CONTROL evaluation.

Simulating the BNP Scenarios

Using the MS SEL scenarios as starting point, pre-selection was supposed to be ignored at national level, leading to biased national proofs. Bias in (G)EBV was introduced at national level using the results from the study of Patry and Ducrocq (2011b) as a realistic measure of the bias distribution. In that study, a genomic pre-selection step was simulated in the cohort of YB with only one crop of daughters. For a trait with 36% heritability and 10% pre-selection rate, the bias in the YB cohort was -0.227 genetic standard deviation on average with a standard deviation of 0.016.

Hence, in the present study the individual bias (Δ_i) was drawn from a random standard normal variable ~N(-0.227, 0.016). This random value was added to each actual national breeding value N-EBV$_i$ and the sum N-EBV$_i$ + Δ_i was considered as the new input for MACE. International genetic evaluations were run based on these biased national breeding values for the cohort of the YB either from A, B or C (BNP-A, BNP-B, BNP-C scenarios) or all countries at the same time (BNP-all).

Implementation

Scenarios were replicated 10 times. Hence, 10 different lists of sires to evaluate (RD SEL scenarios) and then 10 different lists of national breeding values (BNP scenarios) were assumed to be submitted to international evaluations. The CONTROL and the MS SEL scenarios were not replicated as no random sampling was involved. A total of 84 MACE runs were performed and compared to the CONTROL run. Bias was defined as the mean difference between MACE EBV from simulated scenarios and MACE EBV from the CONTROL scenario, averaged over all replicates. To assess the impact of bias on selection decisions, the country of origin and the age group of the top 100 sires were identified on each country scale and for each scenario and compared to the CONTROL situation.

RESULTS

Consequences of Genomic Selection on MS and MACE EBV Distributions

Genomic selection was mimicked among YB by retaining the bulls with the highest MS estimate within large half-sib families in each of the 3 countries. Table 3 describes on each country scale the average MS estimate and MACE EBV before selection (all candidates were considered) and after a random or MS selection. As expected, genomic selection led to higher and less variable MS and MACE EBV. The hypothesis on MS distribution underlying the mixed model methodology was clearly violated, i.e. the MS mean was significantly different from zero on all scales, and its standard deviation was reduced compared to the CONTROL case. When young bulls were selected at random (RD SEL), only the quantity of information involved in MACE evaluations was affected. In such cases, MS deviation was found to be not different from zero and MACE EBV were not biased.

Effect on International Evaluations of including Pre-selected Proofs from one Country

This corresponds to the MS SEL-A, MS SEL-B and MS SEL-C scenarios. For illustration, Figure 1 presents the distribution of bias on each country scale when pre-selection took place in country A (scenario MS SEL-A). Distribution of bias for scenarios MS SEL-B and -C are presented in additional files (1, 2). Table 4 displays bias among the foreign YB on the scale of the country submitting biased proofs. MACE breeding values of foreign YB were overestimated (Table 4) on the local scale. Table 5 displays bias among domestic YB on all scales. It appears that in all cases MACE breeding values of domestic YB on their own scale were virtually unbiased (diagonal values in Table 5). In contrast, MACE breeding values of the latter were clearly biased downward on foreign scales (off-diagonal values in Table 5). These trends were observed whatever the country submitting pre-selected domestic proofs. Expressed in genetic standard deviation, bias among domestic YB (Table 5) ranged from -0.10 to -0.33 on foreign scales whereas bias among foreign YB (Table

4) ranged from 0.07 to 0.17. Among the list of top 100 sires, the proportion of YB actually increased from 5 to 11% (Table 6). However, YB from the country providing selected data tended to be replaced by foreign YB.

Combining Effects of Proof Pre-selection with Biased Proofs in one Country

Results from the BNP-A, BNP-B and BNP-C scenarios are shown in Figure 2 and Tables 7 and 8. Figure 2 presents the distribution of bias on each country scale if country A sent partial and biased domestic proofs to Interbull (scenario BNP-A). Distribution of bias for scenarios BNP-B and -C are presented in additional files (3, 4). Table 7 and Table 8 are equivalent to Table 4 and Table 5, but for the BNP scenarios. The observed bias actually tended to be the sum of both effects, pre-selection and biased national proofs. On local scales, the magnitude of bias of MACE breeding values of domestic YB was the same as the magnitude of bias at national level due to genomic selection as if pre-selection had no extra effect. As far as foreign YB were concerned (Table 7), the upward bias was slightly buffered by the negative bias among national proofs. On foreign scales, MACE breeding values of domestic YB were even more underestimated (off-diagonal elements of Table 8). Effect of bias due to genomic selection was however buffered by genetic correlations below 1 on foreign scales. Domestic YB were the most penalized: MACE breeding values were biased downward on all scales, bias ranging from -0.23 to -0.47 in genetic standard deviation (Table 8). It follows that the proportion of domestic YB among the top 100 sires clearly decreased (Table 9), in favour of foreign YB. Note that the magnitude of bias increased when the level of country heritability decreased (from 0.48 to 0.30) and when the number of YB with biased proofs increased (from 131 in B to 360 in C for the same level of heritability).

Effect on International Evaluations of including Pre-selected Proofs from all Countries

We assumed that A, B and C all implemented genomic selection among young bulls, all accounting for it at national level. Therefore, pre-selected but unbiased data were available at Interbull for all countries (MS SEL-all scenario). Looking at domestic MACE breeding values of YB, they were unbiased when expressed on their local (Figure 3). However, all MACE breeding values of foreign YB were systematically affected: depending on the country of origin, they were either underestimated or overestimated on the same foreign scale. For example, MACE breeding values were overestimated for B YB, while they were underestimated for C YB on A scale (Figure 3, on the left). The bias ranged from - 0.07 to 0.06 genetic standard deviation. This bias barely changed the global proportion of YB among the top 100 sires (Table 10), but few changes in their country of origin still occurred, depending on the sign of the bias.

Combining Effects of Proof Pre-selection with Biased Proofs in all Countries

Results for scenarios BNP-all are presented here. It was assumed that A, B and C implemented genomic selection among young bulls and none of them accounted for pre-selection at national level. Biases observed in MS-all scenario were pushed downward according to the direction and magnitude of bias in national proofs. Bias was the least variable for domestic YB on their local scale (Figure 4). All YB were

highly penalized. Consequently, the proportion of YB among the top 100 sires decreased significantly: 9 to 15 YB, i.e. one quarter of the initial set of YB were removed from the top 100 list in favour of older sires (Table 11). YB from the country submitting the largest number of biased data, i.e. C in these simulations, were the most impacted.

DISCUSSION

Genomic selection generates data with non random missing values and impacts the quality of national BLUP evaluations (Patry and Ducrocq, 2011b). At international level, the issue of missing values but also of biased values occurs in a multi-trait framework. It may involve one, several or all countries and these effects interact with each other. Pre-selection was shown to lead to the violation of the prior assumptions of the MS distribution. Therefore, it was expected that MACE breeding values of the genomically selected YB would be biased in the same way as national breeding values. In the simplest case, when only one country is assumed to submit partial but correct data to Interbull, these YB were indeed always penalized. Their MACE breeding values were clearly underestimated on foreign scales. However, thanks to the correction implemented on domestic evaluations, they were virtually unbiased on their local scale, but MACE breeding values of foreign YB were overestimated. For the same reason, MACE breeding values for domestic YB were always unbiased on their own scale when all countries supply incomplete but correct data for international evaluations. Then, MACE breeding values for foreign YB were either overestimated or underestimated. In multi-trait analyses, the weight given to domestic proofs is much higher than foreign parents and sibs information: this is why the estimation of breeding values for domestic YB appears to be robust on their own scale. However, the comparison of breeding values of bulls from different countries remains relative and the sign of bias depends on the scale of evaluation. In contrast, the magnitude of bias depends on the heritability of the trait and country.

When genomic pre-selection is not accounted for at the national level, bias in national proofs is transmitted into MACE breeding values on the domestic scale. This bias is then transferred to foreign YB through the international genetic relationships. Bias increases with the size of the YB cohort but is buffered by genetic correlations lower than 1. Moreover, this bias tends to be added at international level to the error in breeding values estimation due to pre-selection. Then, overestimations are alleviated whereas underestimations are emphasized. Nevertheless, over- or under-estimation of domestic or foreign YB breeding values may favour or penalize their contribution to their top 100 sires list. YB from countries submitting selected or biased proofs were the most penalized and affected the composition of the top 100 sires: countries submitting incorrect data to Interbull may lose market shares. For importing countries without own genomic evaluation, this may lead to the import of overestimated YB.

In our study, candidates were retained based on their MS value to mimic genomic pre-selection. This is equivalent to a within family selection. Furthermore, given the data available which were real EBV, the MS estimates were more precise than usually expected with genomic evaluation at birth. Hence, the applied selection differential was stronger than actually expected. These assumptions lead to simulations which may be considered as unrealistic. The extent of the observed bias in our simulations must be considered with caution, but its direction and the influence of

the different factors influencing it are not affected by this overestimation, as illustrated for example when selection intensity is reduced (results not shown).

This study showed that each incomplete or incorrect data from genomic selection may cause bias at international level. However, the bias magnitude might be very difficult to predict with the increasing number of countries implementing genomic selection and with the different levels of selection intensity that can be implemented. Strategies regarding genomic selection can be very diverse and may lead to contrasted directions of bias. However, it is certain that missing data and biased data lead to incorrect breeding values and non optimal rankings, likely to impact selection decisions and market shares.

From a technical or a business point of view, all countries participating in Interbull might be affected by practices observed worldwide. Therefore, it is essential to prevent from such a disorganisation.

First of all, it is strongly recommended to account for genomic selection at national level and thus avoid a penalization of YB at national and international levels. The basic requirement is to include all information about selection decisions in national evaluations, i.e. from culled as well as selected candidates. Three procedures have been proposed. One is based on a single-step evaluation procedure (1) considering a relationship matrix that blends full pedigree and genomic information to simultaneously evaluate genotyped and non-genotyped animals(Aguilar et al., 2010, Christensen and Lund, 2010). A second type of approach is to include genotyped but culled candidates into BLUP evaluations, considering de-regressed GEBV as performances (2) as proposed by Ducrocq and Liu (2009) or GEBV as a correlated trait (3) as proposed by Mäntysaari and Strandén (2010). The approach of Ducrocq and Liu (2009) was implemented by Patry and Ducrocq (2011a) for all candidates undergoing genomic selection and bias due to genomic pre-selection was shown to be removed.

However, the breeding values available for culled candidates are based on genomic information (GEBV from one-step or multi-step procedures). Consequently, GEBV should be combined with the national BLUP evaluations into MACE evaluations. The GMACE methodology, a modified MACE for genomics (VanRaden and Sullivan, 2010) aimed at combining N-EBV and GEBV and its principle could also be adapted to such issue. Nevertheless, the number of genotyped animals and thus of genomic breeding values to combine into MACE may dramatically increase. Managing such a huge quantity of data and validating them may become an issue at the Interbull level. Furthermore, this study illustrates that bias trends will be very difficult to predict in a real framework. Consequently, it will be very difficult to detect bias due to genomic selection through any validation procedure. Finally, another problem which was already raised at national level will also exist at international level: if GEBV are combined with N-EBV into MACE, genomic information will be transmitted from genotyped animals to non genotyped animals through the international relationship matrix and be included in the resulting MACE breeding values. Consequently, these latter will no longer be independent from genomic information. However, international reference population require MACE breeding values to define the future genomic predictions. It is feared that genomic information will be double counted and therefore reduce the quality of the future genomic evaluations, as described by Patry and Ducrocq (2011a). To alleviate international genetic evaluations from the constraints of the huge dataset and of double counted

genomic information, another alternative would be to define and estimate effect for genetic groups as applied for base population (Quaas, 1988) but for contemporary animals undergoing genomic selection. Or, an international approach more robust to over- or under-estimation of genetic trends could be developed as previously proposed by Ducrocq et al. (2003).

CONCLUSIONS

This study assessed the bias at international level due to genomic selection. It was shown that using selected and possibly biased national proofs affected the quality of international genetic evaluations delivered by Interbull. The bulls which underwent genomic selection were the most penalized ones. Moreover, biases were transmitted to foreign young bulls on the different scales of evaluations, impacting international ranking and market shares. Consequences on international genetic evaluations might be difficult to predict in complex and heterogeneous situations considering the diversity of breeding practices and policies between countries and within country. Accounting for genomic selection at national level is of high relevance as well as the transmission to Interbull of all available information. The MACE methodology should be adapted to the use of all types of breeding values, including genomic information.

ACKNOWLEDGMENTS

We gratefully thank Interbull for allowing the use of national proofs and for making available tools to implement the MACE methodology for this study. We also thank the Department of Animal Breeding and Genetics at SLU, the PROTEJE and the AMASGEN working groups for their useful remarks. Financing of the AMASGEN project (Jouy-en-Josas, France) by Agence Nationale de la Recherche and APISGENE is gratefully acknowledged.

REFERENCES

Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J Dairy Sci 93(2):743-752.

Christensen, O. and M. Lund. 2010. Genomic prediction when some animals are not genotyped. Genet Sel Evol 42(1):2.

Ducrocq, V., I. Delaunay, D. Boichard, and S. Mattalia. 2003. A general approach for international genetic evaluations robust to inconsistencies of genetic trends in national evaluations. Interbull bulletin 30:101-111.

Ducrocq, V. and Z. Liu. 2009. Combining genomic and classical information in national BLUP evaluations. Interbull Bulletin 40:172-177.

Henderson, C. R. 1975. Best Linear Unbiased Estimation and Prediction under a selection model. Biometrics 31(2):25.

Interbull. 2008. Interbull code of practices Method of international evaluation.

Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J Dairy Sci 92:4656 - 4663.

Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. J Dairy Sci 92(9):4648-4655.

Mäntysaari, E. A. and I. Strandén. 2010. Use of bivariate EBV-DGV model to combine genomic and conventional breeding value estimations. in Proc. 9th WCGALP, Leipzig, Germany.

Nilforooshan, M. A., B. Zumbach, J. Jakobsen, A. Loberg, H. Jorjani, and J. Dürr. 2010. Validation of National Genomic Evaluations. Interbull meeting 42:56-61.

Patry, C. and V. Ducrocq. 2011a. Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. Genet Sel Evol 43(1):30.

Patry, C. and V. Ducrocq. 2011b. Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. J Dairy Sci 94(2):1011-1020.

Pollak, E. J., J. van der Werf, and R. L. Quaas. 1984. Selection bias and multiple trait evaluation. J Dairy Sci 67:1590-1595.

Quaas, R. L. 1988. Additive Genetic Model with Groups and Relationships. J Dairy Sci 71(5):1338-1345.

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

VanRaden, P. and P. Sullivan. 2010. International genomic evaluation methods for dairy cattle. Genet Sel Evol 42(1):7.

Tables

Table 1- Heritability, genetic correlations between countries below the diagonal, and estimated sire standard deviation on the diagonal for each of the 3 countries sending protein yield data to Interbull

Heritability	Country	A	B	C
0.48	A	8.600		
0.30	B	0.851	9.165	
0.30	C	0.870	0.898	9.446

Table 2– Contrasts between simulated scenarios

Scenarios	Pre-selection	Biased national breeding values	MACE runs
CONTROL	No	No	1
MS SEL[1]	Yes (genomic)	No	1 x 4
RD SEL[2]	Yes (random)	No	10 x 4
BNP[3]	Yes (genomic)	Yes	10 x 4

[1] MS SEL-A, -B, -C, -all: 10% of young bulls retained from either A, or B or C or in the 3 countries simultaneously

[2] RD SEL-A, -B, -C, -all: 10% of young bulls retained from either A, or B or C or in the 3 countries simultaneously

[3] BNP-A, -B, -C, -all: 10% of young bulls retained from either A, or B or C country or in the 3 countries simultaneously

Table 3: Mendelian sampling distribution, average genetic level on the 3 scales among young bulls from A, B and C depending on the implemented selection

YB[1] from	Average MS[2] term			Average MACE EBV[3]		
	CONTROL	RD SEL[4]	MS SEL[5]	CONTROL	RD SEL[4]	MS SEL[5]
A	0.00 +/- 0.72	0.00 +/- 0.72	1.22 +/- 0.40	0.97 +/- 0.96	0.99 +/- 0.89	2.20 +/- 0.71
	n=2,234	n=224	n=224	n=2,234	n=224	n=224
B	0.00 +/- 0.71	0.00 +/- 0.71	1.20 +/- 0.36	0.52 +/- 0.92	0.54 +/- 0.86	1.73 +/- 0.70
	n=1,282	n=131	n=131	n=1,282	n=131	n=131
C	0.00 +/- 0.74	0.00 +/- 0.74	1.26 +/- 0.39	0.76 +/- 0.94	0.79 +/- 0.87	2.04 +/- 0.68
	n=3,602	n=362	n=362	n=3,602	n=362	n=362

[1] YB: young bulls

[2] MS: Mendelian sampling

[3] MACE EBV: breeding values estimated by the MACE methodology

[4] 10% random selection among young bulls

[5] 10% genomic selection among young bulls

Table 4: Bias in MACE results expressed on A, B or C scale when one country submitted pre-selected data - among foreign young bulls (scenarios MS SEL-A, -B or -C)

		Country submitting selected proofs		
		A	B	C
		A scale	B scale	C scale
Country of origin	A		0.09	0.11
	B	0.15		0.07
	C	0.17	0.07	

Table 5: Bias in MACE results when one country submitted pre-selected data - among domestic young bulls (scenarios MS SEL-A, -B or-C)

Country of origin	A scale	B scale	C scale
A	-0.01	-0.11	-0.10
B	-0.19	-0.03	-0.13
C	-0.33	-0.28	-0.03

Table 6: Changes on local scale in the proportion of young bulls by country of origin when one country submitted pre-selected data (proportion of young bulls in the CONTROL situation are in parenthesis)

Scenario	Scales	Among the top 100 sires	By country of origin:		
			A	B	C
MS SEL-A	A	+ 11 (54)	-2	3	10
MS SEL-B	B	+ 8 (53)	4	-4	8
MS SEL-C	C	+ 5 (59)	9	2	-6

Table 7: Bias in MACE results when one country submitted pre-selected and biased data - among foreign young bulls (scenarios BNP-A, -B or -C)

		Country submitting selected proofs		
		A	B	C
		A scale	B scale	C scale
Country of origin	A		0.07	0.09
	B	0.12		0.06
	C	0.14	0.05	

Table 8: Bias in MACE results when one country submitted pre-selected and biased data - among domestic young bulls (scenarios BNP-A, -B or -C)

Country of origin	A scale	B scale	C scale
A	-0.23	-0.28	-0.28
B	-0.35	-0.25	-0.31
C	-0.47	-0.44	-0.25

Table 9: Changes on local scale in the proportion of young bulls by country of origin when one country submitted pre-selected and biased data (scenarios BNP-A, -B or –C) - proportion of young bulls in the CONTROL situation are in parenthesis

Scenario	Scales	Among top 100 sires	By country of origin:		
			A	B	C
BNP-A	A	+ 3 (54)	-10	3	10
BNP-B	B	+ 3 (53)	4	-10	9
BNP-C	C	- 2 (59)	8	4	-14

Table 10: Changes in the proportion of young bulls by country of origin when the 3 countries submitted pre-selected data (scenario MS SEL-all) - proportion of young bulls in the CONTROL situation are in parenthesis

Scenario	Scales	Among top 100 sires	By country of origin:		
			A	B	C
MS SEL-all	A	0 (54)	-1	2	-1
MS SEL-all	B	0 (53)	2	0	-2
MS SEL-all	C	+ 3 (59)	3	2	-2

Table 11: Changes on local scale in the proportion of young bulls by country of origin when the 3 countries submitted pre-selected and biased data (scenario BNP-all) - proportion of young bulls in the CONTROL situation are in parenthesis

Scenario	Scales	Among top 100 sires	By country of origin:		
			A	B	C
BNP-all	A	- 9 (54)	-4	1	-6
BNP-all	B	- 15 (53)	0	-7	-8
BNP-all	C	- 12 (59)	-1	-1	-9

Figures

Figure 1: Bias distribution in MACE results on the three scales among young bulls by country of origin when country A submitted pre-selected data (scenario MS SEL-A)

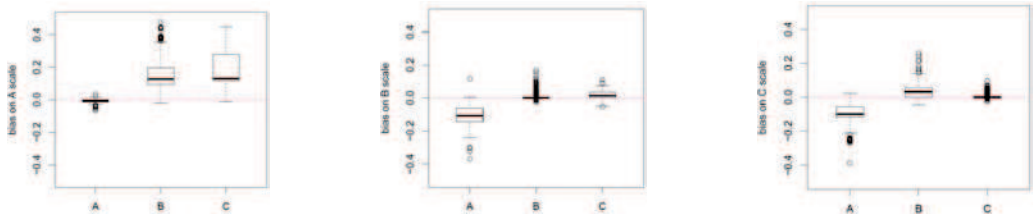


Figure 2: Bias distribution in MACE results on the three scales among young bulls by country of origin when country A submitted pre-selected and biased data (scenario BNP-A)

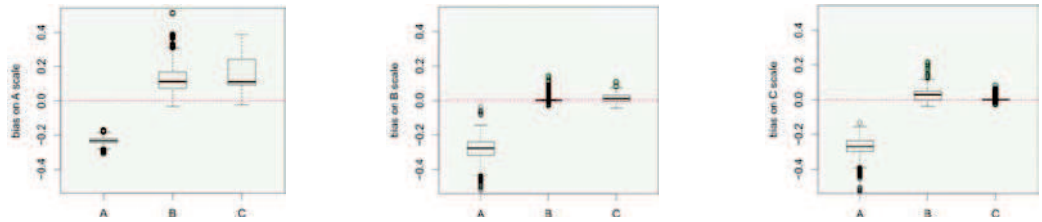


Figure 3: Bias distribution in MACE results on the three scales among young bulls by country of origin when country A, B and C submitted pre-selected data (scenario MS SEL-all)

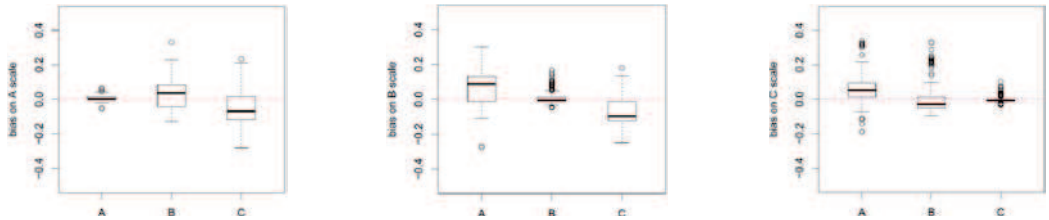
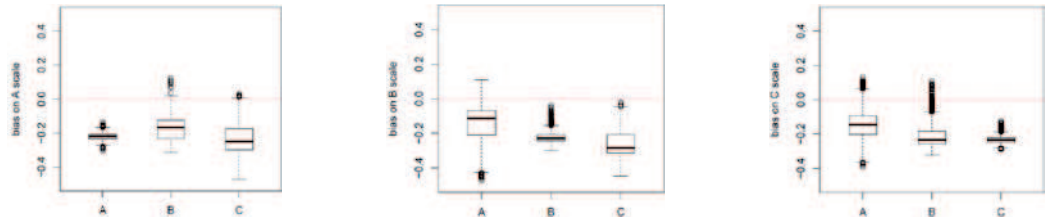
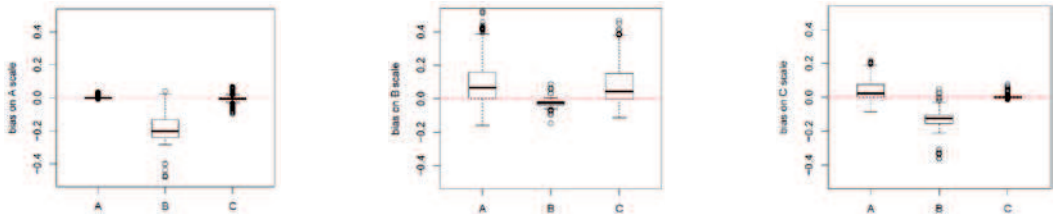


Figure 4: Bias distribution in MACE results on the three scales among young bulls by country of origin when A, B and C submitted pre-selected and biased data (scenario BNP-all)

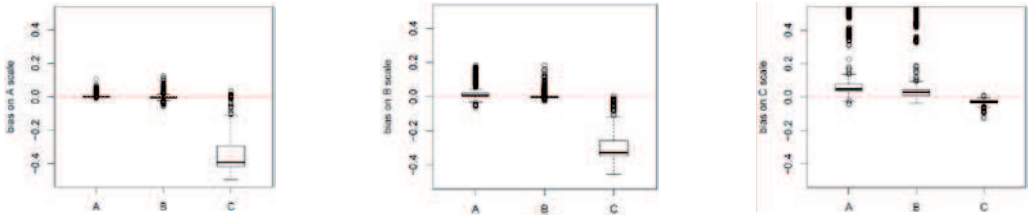


APPENDIX

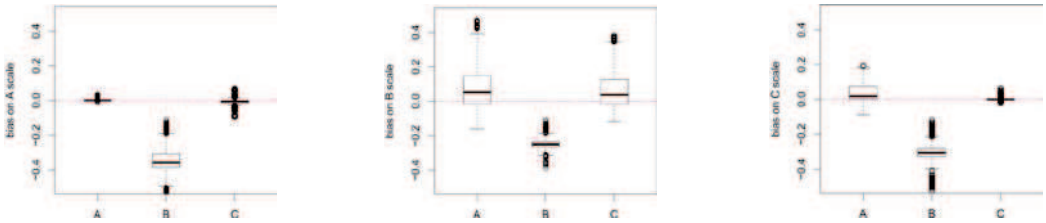
Additional file 1 – Bias distribution in MACE results on the three scales among young bulls by country of origin when country B submitted pre-selected data (scenario MS SEL-B)



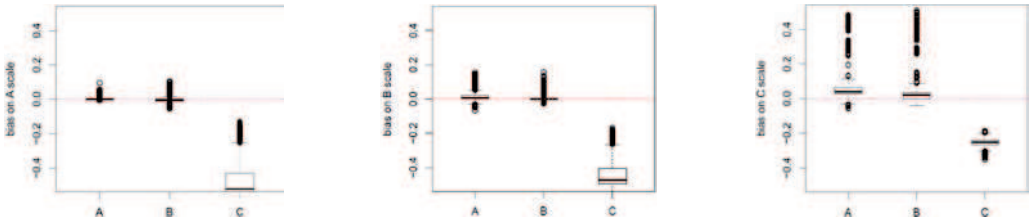
Additional file 2 – Bias distribution in MACE results on the three scales among young bulls by country of origin when country C submitted pre-selected data (scenario MS SEL-C)



Additional file 3 – Bias distribution in MACE results on the three scales among young bulls by country of origin when country B submitted pre-selected and biased data (scenario BNP-B)



Additional file 4 – Bias distribution in MACE results on the three scales among young bulls by country of origin when country C submitted pre-selected and biased data (scenario BNP-C)



CHAPTER 7 - General Discussion

7.1. Limits and contributions of the study

7-1.1. LIMITS OF THE SIMULATION STRATEGIES AND ANALYSES

- **A simplified design for simulations**

In articles I and III, bias was measured in national and international genetic evaluations after genomic selection. Assessments were occasionally based on simplistic assumptions and simplified designs of breeding schemes. Genomic selection was assumed to be only implemented on one trait. However, the chosen strategy involved repeated simulations to remove random errors. Several factors influencing bias in genetic evaluations were studied under controlled conditions. In the first article, selection based on molecular information was studied for different levels of heritability, of precision of GEBV and of proportion of missing data. In the third article, genomic selection was studied in scenarios involving two potential sources of bias related to the use of incomplete (pre-selected) and possibly incorrect (biased national proofs) data sets in international evaluations. Single trait evaluation (article I) and multi-trait evaluation (article III) were both used. The effect of correlations between traits was studied in article III. In fact, the effects of the different sources of bias or of the various factors influencing the bias magnitude were studied one at a time before simulating more complex and more realistic scenarios (article III).

Although far from the complexity and heterogeneity of real breeding schemes, such designs have the advantage of isolation of each source of bias as well as each factor influencing its magnitude to facilitate interpretation.

- **Lack of actual GEBV to mimic genomic selection and to measure bias**

Liu *et al.* (2009) also measured bias in classical evaluations due to genomic selection. Unlike them, real DGV or GEBV were not used here to mimic genomic selection or to measure bias at either the national (article I) or the international level (article III). Only progeny tested bulls with known pedigree, performances and EBV were considered as a reference to assess bias. The complete data set was used to simulate TBV and GEBV, or altered to mimic the effect of genomic selection among candidates. Such strategies benefited from the use of actual data and accounted for the effects of the large population size, real pedigree and data structure.

However, these strategies are not completely realistic. In article III, genomic selection was approximated as a selection on MS deviations within family, i.e., among half-sibs, and country. In such a case, MS deviations were computed from quite reliable international breeding values. This choice probably leads to an overestimation of the selection response compared with a selection based on actual GEBV given the same selection intensity. Selection was also implemented within family, i.e., among full-sibs, in article I. This was a

selection on MS deviation as well but based on a lower accuracy. In both cases (articles I and III), ignoring family information for selection decisions may overestimate bias.

- **A restricted set of bias indicators**

Expressed in genetic standard deviations, bias was measured in two ways:

- In article I, true breeding values were assumed known and bias was measured as the average difference between true and estimated breeding values.
- In article III, true breeding values were not simulated and bias was assessed comparing estimated breeding values obtained from altered or unaltered data sets. This was also implemented by Liu *et al.* (2009) to measure bias due to genomic selection: they compared proofs from altered data set and proofs from reference evaluations, both expressed in percent of genetic standard deviation.

Other bias indicators could also have been used. For example, it is common to study the linear regression of TBV on EBV or of EBV from unaltered data set on EBV from altered data set. Only based on average differences but under various conditions, the same trends were observed though, i.e., a penalization of the genomically selected sires. This increases confidence in our results.

Liu *et al.* (2009) also studied the correlation between both sets of proofs. Such an indicator was not presented in article III but was very convenient in article I as it also provided true reliability.

The choice of the presented indicators was motivated by the will to make them easy to understand, and to avoid confusion.

7-1.2. A BETTER UNDERSTANDING OF THE BIAS MECHANISMS AFTER SIMULATIONS OF GENOMIC SELECTION

- **The characteristics of genomic selection and their simulations**

From the literature review on biases (Section 3-3.1), genomic selection has been associated with two features that BLUP does not support: sequential selection and selective phenotyping.

On the one hand, the selection criterion, i.e., GEBV, is actually absent from the evaluation model despite the fact that it is obviously correlated to the breeding value of the trait of interest. In article I, GEBV were jointly distributed with TBV and EBV and selection was based on the GEBV but the evaluation model did not include this knowledge of GEBV.

On the other hand, genomic selection generates a data subset which is not representative of the entire population of candidates. In articles I and III, missing data was generated by deleting phenotypes.

Both features were simulated in our studies and led together to a selection process which is not fully described by the genetic evaluation model. The null expectation of MS was no longer true. This was checked in the three articles among young sires being evaluated. This

MS indicator measures how large the deviation from the initial assumption is and was used as an indicator to validate that the genomic selection step was correctly implemented in our simulations.

The results of the simulations confirmed that violating BLUP assumptions decreases the quality of the predictions.

- **Key factors in the generation of bias**

From our simulation results the proportion of selected animals (10% or 25%), the heritability of the trait (4 levels between 0.14 and 0.48), the precision of genomic evaluations (27%, 50%) or the precision of MS estimates were shown to influence the magnitude of the bias in breeding value estimations. Reducing any of these factors simply increases the inconsistency between true and estimated MS means or variances.

In fact, when the proportion of culled candidates increases, the number of missing animals in the relationship matrix also increases. Moreover, the deviation from a zero MS mean is larger: the bias among the selected candidates increases, as noticed by Sorensen and Kennedy (1984). In their study on selection bias for three different culling rates, Liu and al. (2009) also noticed that the number of animals with significantly biased breeding values was higher when culling rate increased.

The difference between observed and assumed distributions of breeding values changes with the variance ratio, i.e., with the trait heritability: a lower heritability puts more emphasis on any inconsistency and increases the bias magnitude. Sorensen and Kennedy (1984) also reported the role of heritability in the bias magnitude.

The quantity of genomic information in genomic predictions was summarized by the level of genomic effective daughter contributions (gEDC). In article I, for the same level of heritability, two different levels of gEDC were tested. When genomic predictions are more accurate, bias is less important.

In article III, the genetic correlation between the three countries differed: higher genetic correlations between country and more generally traits generated a larger transfer of the bias from one trait to another. This was also reported by Pollak *et al.* (1984).

In conclusion, one of the main assets of the chosen strategy is to provide a qualitative assessment of bias under controlled, potentially unrealistic, but diversified conditions. Such an approach must be seen as a case study to figure out the possible impacts of genomic selection on the quality of genetic evaluations when the intensity of genomic selection is changing. However, methods of this type cannot be implemented to routinely measure bias in classical evaluations.

7.2. Consequences on classical and genomic evaluations

7-2.1. THE EXTENT OF BIAS IN CLASSICAL EVALUATIONS

- **A wide propagation of bias among cohorts, traits, and generations**

Estimates of breeding values were differentially affected according to the cohorts of animals considered in the population to evaluate.

➢ Among selected young bulls with daughters:

In our simulations, national BLUP breeding values were significantly underestimated when the genomic selection step was not accounted for. Bias reached one quarter to one third of a genetic standard deviation at the national level. The same trend was also observed in MACE evaluations when breeding values of the selected young sires were expressed on foreign scales. When MACE breeding values were expressed on the scale of the country which implemented genomic selection of young sires, these bulls were also penalized: their breeding values were not biased but MACE breeding values of contemporary foreign sires were significantly overestimated. Hence, bulls undergoing genomic selection were systematically penalized.

➢ Among progeny:

Bias due to genomic selection is transferred to relatives of genomically selected sires through the genetic relationship matrix in BLUP or MACE methods. In article I, about half the bias was transferred from selected young to their progeny.

➢ Among sires and dams of the young bulls:

A side effect of genomic selection is that the information that should have been included in the model is now missing. Sires and especially dams of culled bulls are also affected (Liu *et al.*, 2009). In their study, EBV of sires and dams of the culled genotyped bulls were overestimated. Bias was higher for sires with all of their sons eliminated from the evaluations and bias was much higher for dams because the dams had no records of their own in these simulations and their EBV was strongly influenced by their genotyped sons.

Because of genetic correlations between countries, bias due to genomic selection is also transferred to foreign populations in a multi-trait analysis. Article III showed that bias was transferred on different scales at international level. It can be deducted that bias will also be transferred between correlated traits at the national level when multi-trait models are implemented for genetic evaluations. Total merit indices may be affected.

Furthermore, new bias will be in each generation created with the new cohorts of candidates for genomic selection. Impacts on estimation of genetic parameter estimates were not studied. From the literature review on bias in genetic evaluations (Section 5-1.2), there is a fear that the accumulation of biased breeding values over generations would also lead to biases in the estimation of genetic parameters participating to long-term problems.

- **Facing more complex conditions: an unpredictable magnitude and direction of bias**

In a more realistic framework, the key factors influencing bias (Section 7-1.2) interact with each other so that it becomes difficult to predict bias magnitude and direction, e.g, whether bias corresponds to an under or overestimation of breeding values.

For example, genomic selection can be applied simultaneously on many traits with positive or negative correlations between traits. Article III showed that the direction of bias was difficult to predict when a multi-trait model was used because of the different weights given to information sources in different countries' scales. Furthermore, these traits can have different levels of heritability more or less accentuating bias magnitude. The proportion of retained candidates can vary a lot given the diversity of breeding schemes and of selection objectives within and between countries. Moreover, genomic predictions are associated with different levels of accuracy depending on the size of the reference population (RP), the quality of the genomic prediction model and method, the chip density and the quality of imputation methods among other factors. Genomic selection can occur among male and female cohorts over several generations. The situation will become very complex when daughters of genomically selected sires or genomically selected dams are mated to genomically selected sires. As all these factors influence bias in opposite directions and intensity, bias might widely propagate in an uncontrolled way.

- **Persistence of bias despite daughter contribution**

In the past, another classical source of bias in national genetic values has been data from daughters born from imported semen, i.e., from foreign and highly selected sires (Bonaiti and Boichard, 1995). In this situation, bias tended to be reduced when an increasing number of daughters of these sires had their performances included in the national evaluation system (Pedersen *et al.*, 1999). But in the case of genomic selection, such bias remains:

- the EBV of these daughters ($\widehat{a_{dau}}$) will be first affected by half the bias (Δ) transferred by their sires (s) through the relationship matrix (article I):

$$\widehat{a_{dau}} = \frac{1}{2}(\widehat{a_s} + \Delta) + \frac{1}{2}\widehat{a_d} + \widehat{\varphi_{dau}} \Leftrightarrow \widehat{a_{dau}} = (\frac{1}{2}(\widehat{a_s} + \widehat{a_d}) + \widehat{\varphi_{dau}}) + \frac{1}{2}\Delta \quad [26]$$

with $\widehat{a_d}$ the dam breeding value and $\hat{\varphi}_{dau}$, the estimated MS term.

- the increasing number of daughters will raise the accuracy of the sire EBV and especially of the MS estimates. The MS deviation from null expectation is therefore more certain.

- **A worse situation: what about data recording after abandon of progeny testing?**

It is also feared that disuse of progeny testing may disorganize the existing regulated and widespread performance recording. This may enlarge biases in genetic evaluations if data records are no longer a representative sample of high quality data. So far, the importance of data recording is still widely accepted even with genomic selection but a segmentation of the herd population may occur (Ducrocq and Santus, 2011) with different levels of recording in quality and quantity: herds with high quality and exhaustive recording, commercial herds with simplified recording and research herds with experimental recording on new traits.

7-2.2. WIDESPREAD PROBLEMS FROM BIASED CLASSICAL EVALUATIONS

- **Impact on genomic evaluations**

In reference populations, by-products of EBV, e.g., DYD or DRP, from classical evaluations are used as phenotypes and participate to the estimation of molecular marker effects or of breeding values based on genomic matrices. In Holstein, thousands of bulls already progeny tested were genotyped (de Roos, 2011) to set up the RP. These EBV came from the national or international genetic evaluation systems after progeny testing and were assumed to be correct. If genomic selection is not accounted for in classical evaluations, the bias observed in national or international EBV will also affect the definition of the future genomic predictions. Such a bias was not yet measured but using biased data as an input, it is certain that the quality of the genomic evaluations will be affected too. The role of classical EBV in the definition of equations for genomic evaluation highlights the importance of maintaining unbiased BLUP and MACE evaluations.

- **Impact on countries without genomic evaluations**

The danger after genomic selection is that this bias will propagate from biased and selected national proofs to the international level. As a result, MACE breeding values would be biased not only on the scale of the countries implementing genomic selection but also on the scale of importing countries with no genomic evaluation. This may lead to the import of genomically selected young bulls based on underestimated breeding values as well as import of bulls from other cohorts or countries based on overestimated breeding values, as shown in article III. Import decisions would no longer rely on fair indicators. Such a bias would first penalize the country with genomic selection but it would also lower the breeding efficiency of the other countries.

- **Altered rankings and impacts on semen trade**

From the breeder's point of view, breeding values are only relative and are computed to be compared among bulls for selection decisions. It is important to figure out the real impact of the biases on national and international comparisons.

Impact on re-ranking of young sires among the international top 100 sires was examined in article III. Young sires selected based on their GEBV tended to be re-ranked and penalized overall. Older sires or foreign contemporary sires were favored.

One can envision that selection decisions will be mainly based on GEBV in the future. However, it was shown that GEBV could be also biased due to biased classical EBV. Such a re-ranking may also occur with GEBV and definitely affect selection decisions.

AI companies usually compare top list rankings on different scales. Re-rankings may reveal inconsistencies between different evaluation systems or scales. If they do not trust breeding values of some sires, they tend to avoid these bulls. Such biases in genetic evaluations would certainly imply market losses for exporting countries.

- **Impact on selection efficiency**

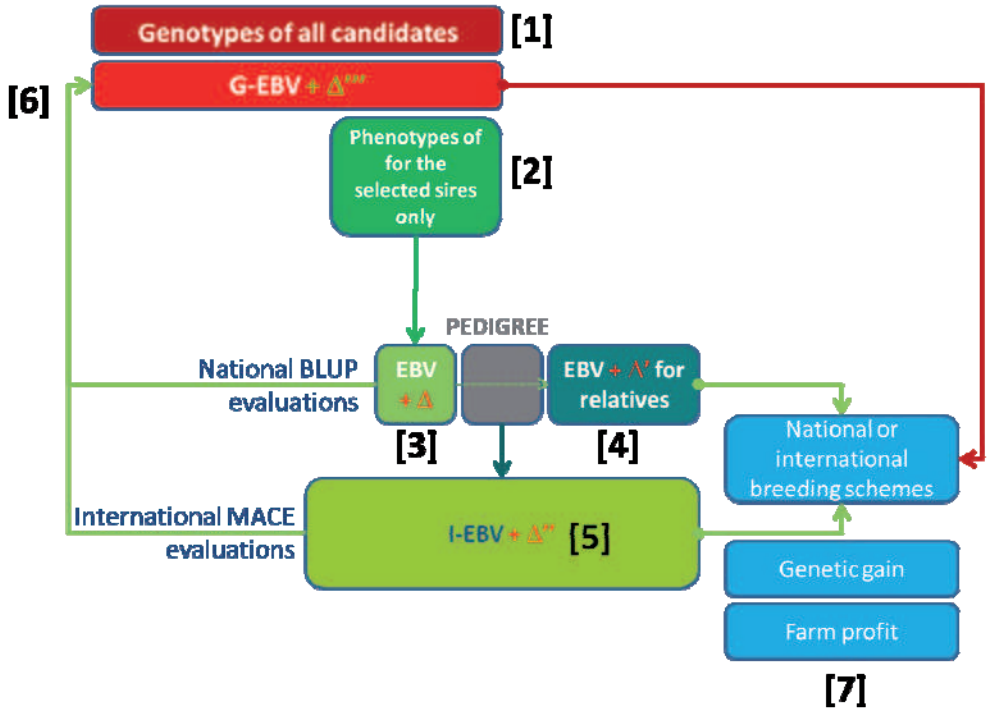
Bias was only studied in classical evaluations over one generation in our study or in the study of Liu *et al.* (2009). From our simulations, the impact on genetic trend cannot be directly analyzed. It is even feared that the estimation of genetic trend might be underestimated but also that its true value might be reduced.

Because selected bulls are underestimated, the selection differentials on the male side will be underestimated too and the estimated genetic trend will be lower than the true genetic trend. Colleau (1989) also reported a discrepancy between true and estimated genetic trends when evaluations based on observations biased by preferential treatment (e.g., BST growth hormone in his case) were used. After a genomic selection step, not only breeding values are biased but also the estimation of genetic trend.

In article I, the increase in MSE and uncertainties among rankings in article III reflected a loss in accuracy of prediction. As a result, selection efficiency will no longer be optimal after a genomic selection step if this step is not accounted for in the classical evaluation and the true genetic trend will be reduced.

Conclusion: **Measures of bias due to a genomic selection step of young sires in one country and for one trait only were shown by simulations to be not negligible. With a likely higher genomic selection intensity in the future, bias is expected to be stronger still. In more real and complex cases, it is difficult to predict its magnitude and direction but it is certain that this bias can propagate to other cohorts and populations: estimated breeding values will be biased first for males or females undergoing genomic selection and then in an uncontrolled way to related animals, genotyped or not, finally to other traits and foreign country scales (as illustrated in Figure 1). Such bias may also accumulate over generations. It would not only affect the national breeding schemes but also the quality of genomic predictions, the international rankings and semen trade. Maintaining unbiased BLUP and MACE evaluations is crucial to optimize the selection efficiency of breeding schemes across the world based on classical and/or on genomic breeding values.**

Figure 1: Propagation of bias throughout the genetic evaluation process after a selection step (2) based on genomic information (1)



Legend: bias is propagating from national EBV on selected young sires with daughters (3) to relatives (4), to international EBV (5), to genomic EBV (6) and its impact on genetic gain (7) according to different magnitude and bias (initially Δ, and then Δ', Δ'', Δ''' after propagation through the different genetic matrices).

7.3. Techniques and methods to avoid bias propagation

It is not possible to ignore the bias caused by genomic selection in classical evaluations. The magnitude and direction of such a bias cannot be predicted for the future. It would be difficult to develop a method to correct bias after propagation in the various populations and traits. To maintain the quality of (inter)national evaluations, it is necessary to:

- **prevent** national evaluations from being biased so that bias is not transferred to non genotyped animals and foreign populations;
- **avoid** bias at the international level. Biases in the estimation of national breeding values and incomplete information should be detected in one way or another to be eliminated from international evaluations.

7-3.1. METHODS TO PREVENT BIAS IN NATIONAL GENETIC EVALUATIONS

To prevent bias, all information on which selection is based should be included in the evaluation model. We will assume first that the list of selection candidates is known and that their pedigree and GEBV are registered in the national databases. In this optimistic case, one straightforward approach is to de-regress GEBV and include the results as weighted genomic pseudo-performances in HMME as (implemented in article II). The relationship matrix should be first complete and correctly computed. In article II, all genomic pseudo-performances were associated with a unique genomic effective daughter contribution (i.e., gEDC). It was also assumed that no sires of candidates were genotyped. This approach was shown to be conceptually satisfactory as bias among young sires and their progeny was removed. However, some issues were not solved and require further studies before actual implementation in routine evaluations. In particular, redundancy of genomic information and dependency between genomic and BLUP evaluations need to be considered.

- Issue 1: the weight given to genomic information as fraction of the genetic variance explained by markers.

Genomic reliabilities tend to be inflated (Van Raden *et al.*, 2009) if one assumes that markers capture all genetic variance. This is not the case because all QTL and neighboring markers are very rarely in complete linkage disequilibrium. Then, the weight given to genomic information, i.e., gEDC derived from genomic reliabilities, is likely to be overestimated.

- Issue 2: the double-counting of genomic information in classical breeding value estimates.

The accuracy of classical evaluations including genomic pseudo-performances might be overestimated too, especially for (genotyped) sires with many genotyped progeny. De-regressed GEBV are included in BLUP evaluations and genomic information is transmitted through the relationship matrix from parents to progeny and vice-versa. However, the total amount of additional information from genotyped relatives cannot exceed the gain in accuracy

from genotyping the animals themselves. Genomic information is especially redundant when both progeny and sires are genotyped.

- Issue 3: the dependency between classical and genomic information by predicting genomic effects from phenotypes already including genomic information.

De-regression of GEBV (or DGV) removes the pedigree information to only include the extra genomic information from the pseudo-performances in national BLUP evaluations. In such a case, the classical evaluations will include genomic information. BLUP solutions of genotyped animals are usually de-regressed to become phenotypes and participate in the estimation of genomic effects. It is essential that these de-regressed proofs are unbiased but it is also important for them not to include undue genomic information. Otherwise, genomic information might be double counted, directly at a molecular level and indirectly at a performances level. This dependency between the two types of evaluations could impact the quality of future genomic predictions (Wiggans, personal communication, 9th WCGALP, Leipzig, 2010).

To deal with all these issues, possible developments are considered hereafter.

- **Computation of individual gEDC**

To implement the approach developed in article II, the computation of gEDC for each candidate sire should be first improved to appropriately account for the correct amount of genomic information in BLUP evaluations and avoid double-counting of genomic information (issue 2). Article II implemented the approach proposed by Ducrocq and Liu (2009) to combine DGV with EBV. They also proposed a method to limit the overestimation of gEDC. Two cases were distinguished depending on the genomic evaluation method.

- Alternative 1: genomic breeding values are computed based on molecular marker estimates.

The correct gEDC can be computed modifying the "information source" method (Harris and Johnson, 1998) to compute reliability. This method consists in progressively combining independent sources of information x and y with reliabilities R^x and R^y to get the global reliability R^{x+y} as $R^{x+y} = \frac{R^x + R^y - 2R^xR^y}{1 - R^xR^y}$. Manipulating this expression, we can also write $R^x = \frac{R^{x+y} - R^y}{R^{x+y}R^y + 1 - 2R^y}$, which is the reliability of x given the reliabilities of y and $x+y$. The information source method is usually implemented in order to gradually combine three independent sources of information coming from the pedigree (*ped*, with reliability $R^2_{i,ped}$), the individual itself (*go*, $R^2_{i,go}$) and the progeny (*prog*, $R^2_{i,prog}$) to get the final individual reliability of animal i (R^2_i).

If we consider only genomic information and if we assume an initial (iteration [0]) value of gEDC= gEDC$^{[0]}$ identical for all animals, we have $R^2_{i,DGV} = R^{2[0]}_{i,go} = \frac{\text{gEDC}}{\text{gEDC} + \alpha}$ where α is the variance ratio. This value could be the genomic daughter equivalents of Van Raden *et al.* (2009). With this starting value, we can compute $R^{2[0]}_i$ using the information source method. But this value $R^{2[0]}_i$ may be incorrect because of double counting: $R^{2[0]}_i$ should not be larger than $R^2_{i,DGV}$ because knowing the genotypes of progeny or parents of *i* does not bring any new genomic information when the genotype of i is already known. The idea developed in Ducrocq and Liu (2009) is to compute $R^{2[k]}_{i,go}$ imposing the constraint $R^{2[k]}_{i,go} \le R^2_{i,DGV}$. The new $R^{2[k]}_{i,go}$ is computed from the current values of $R^{2[k-1]}_{i,ped}$, $R^{2[k-1]}_{i,prog}$ and the expected final result $R^2_{i,DGV}$ using the above formulae. After a few iterations, $R^{2[k]}_{i,go}$ converges to a final value $R^2_{i,go}$ and a $gEDC_i$ specific to animal *i* can be computed as $\text{gEDC}_\text{i} = \alpha \frac{R^2_{i,go}}{1 - R^2_{i,go}}$. This gEDC$_\text{i}$ is used as weight of pseudo-performances. It is assumed to be bounded between 0 and this initial value: $0 \le \text{gEDC}_\text{i} \le \text{gEDC}^{[0]}$.

- Alternative 2: genomic breeding values are estimated from mixed model equations based on a genomic relationship matrix (e.g., using G-BLUP).

The direct genomic reliability, $R^2_{i,DGV}$, can be computed from the difference between the reliability of genomic predictions (evaluation G2 as in chapter 2) and the reliability of BLUP solutions (evaluations C3) when evaluations only include genotyped animals:

$R^2_{i,DGV} = R^2_{i(G2)} - R^2_{i(C3)}$ [27]

This reliability gain is then converted into gEDC as: $gEDC = \frac{\alpha R^2_{i,DGV}}{1 - R^2_{i,DGV}}$ [28].

These methods are based on approximations and may still involve inconsistencies. With the alternative 1 above, gEDC are more satisfying than without adjustments but reliabilities still tend to be overestimated for sires with many genotyped sons. Nevertheless, these methods have the advantage to limit the inflation of reliabilities so that they appear to be more realistic. In Germany (VIT), alternative 2 has been applied to compute gEDC and the associated genomic pseudo-performances (Liu *et al.*, 2009).

- **A global method to distribute genomic information to the whole population: a bivariate analysis**

Genomic breeding values (DGV or GEBV) can be analyzed as pseudo-performances correlated to the actual performances in a multi-trait animal model setting. This principle was already proposed by Gianola *et al.* (2006) (also see Section 2-1.3, case 3). In fact, all candidates are considered in a single model which corrects for sequential selection. Two

bivariate models were recently proposed to further deal with double-counting of genomic information.

- The bivariate model proposed by Mäntysaari and Stränden (2010):

Mäntysaari and Stränden (2010) assumed DGV and DYD to be correlated. Their bivariate model can be written as follows, for bull, *i*:

$$\begin{bmatrix} q_C \\ q_G \end{bmatrix} = \begin{bmatrix} Z_{1i} & 0 \\ 0 & Z_{2i} \end{bmatrix} \begin{bmatrix} a_C \\ a_G \end{bmatrix} + \begin{bmatrix} e_C \\ e_G \end{bmatrix} \quad [29]$$

where Z_{1i} = 1 for sires with daughters, 0 otherwise and Z_{2i} = 1 for genotyped animals, 0 otherwise. q_C refers to the classical performances DYD and q_G to the genomic pseudo-performances, DGV here. Note that they were not de-regressed in contrast with our approach described in article II. a_C refers to the breeding values from classical information and a_G refers to the breeding values from genomic information. As in the classical mixed model equations both "traits" are analyzed using the full relationship matrix **A** and the following matrix of genetic variances:

$$Var \begin{bmatrix} a_C / \sigma_{a_C} \\ a_G / \sigma_{a_G} \end{bmatrix} = \begin{bmatrix} 1 & \rho(q_C, q_G) \\ \rho(q_C, q_G) & 1 \end{bmatrix} = \begin{bmatrix} 1 & \rho(a_C, a_G) \\ \rho(a_C, a_G) & 1 \end{bmatrix} \quad [30]$$

From selection index theory, the authors assumed that: $\rho(a_C, a_G) = R^2_{DGV}$ so that the genetic matrix of variances become after data scaling (*):

$$Var \begin{bmatrix} a_C * \\ a_G * \end{bmatrix} = \begin{bmatrix} 1 & R^2_{DGV} \\ R^2_{DGV} & 1 \end{bmatrix} \quad [31]$$

and $Var \begin{bmatrix} e_C \\ e_G \end{bmatrix} = \begin{bmatrix} \sigma^2_{e_C} / EDC & 0 \\ 0 & \varepsilon \end{bmatrix}$ [32]

with ε very small but different from 0.

Genomic information is transferred from genotyped to non genotyped animals through the genetic correlation between traits and the relationship matrix. The main advantage is that the gain in reliability due to genomic information is naturally bounded by the genetic correlation between the 2 traits, which is even approximated to be R^2_{DGV}. Genomic reliabilities are less overestimated (issue 1) but the assumption on genetic correlation between trait and pseudo-trait is strong.

To deal with the issue of information redundancy (issue 2), Mäntysaari and Stränden considered that if the DYD of a bull is already used for calibration of genomic predictions, information for bulls in the RP should be discounted by decreasing the reliability of their genomic evaluations. This avoids genomic information from bulls in the RP to inflate

reliability of genotyped progeny but it does not consider the inflation of bull reliability due to many genotyped progeny.

- Bivariate model proposed by Stoop *et al.* (2011)

Stoop *et al.* (2011) proposed another multi-trait approach where DGV were analyzed without incorporating the numerator relationship matrix (i.e., replacing it by an identity matrix). The aim is clearly to avoid double-counting of genomic information (issue 2). Genomic pseudo-performances q_G were described using a mass-selection model, i.e., ignoring pedigree information. Estimated genomic breeding values, i.e., DGV, were divided by their reliability, assumed to be identical for all genotyped animals and equal to the trait heritability ($R^2_{i,DGV} = h^2$):

$$q_{i,G} = \widehat{a_{i,G}} / R^2_{i,DGV} \quad [33]$$

The genetic correlation between the trait additive genetic effects under a polygenic and a genomic model were assumed to be 1: $\rho(a_C, a_G) = 1$. In such a case, the genetic variance-covariance matrix is singular. To overcome the resulting problems, Stoop *et al.* (2011) suggest to multiply covariances by a factor β less than 1. In fact, genetic correlations are decreased and the heritability of the genomic effects is increased by β^{-2}. The matrix of variance-covariance can be written:

$$Var\begin{bmatrix} a_C \\ a_G \end{bmatrix} = \begin{bmatrix} \sigma^2_{a_C} & \beta\,\mathrm{cov}(a_G, a_C) \\ \beta\,\mathrm{cov}(a_G, a_C) & \beta^{-2}\sigma^2_{a_G} \end{bmatrix} \quad [34]$$

Assuming, $\beta = R^2_{DGV}$, this becomes equivalent to:

$$Var\begin{bmatrix} a_C \\ a_G \end{bmatrix} = \begin{bmatrix} 1 & R^2_{DGV} \\ R^2_{DGV} & 1 \end{bmatrix} \sigma^2_a \quad [35]$$

which is the genetic variance-covariance matrix used by Mäntysaari and Stränden (2010) (if h^2 is also assumed to be 1).

The main advantage of the method of Stoop et al (2011) over the other one is that it only transfers genomic information between genotyped and non genotyped animals through the relationship matrix of genetic effects and not that of genomic effects. In fact, genomic information still indirectly propagates through the inverse of the genetic relationship matrix.

In all of these methods (Ducrocq and Liu, 2009, Mäntysaari and Strandén, 2010, Stoop *et al.*, 2011), genomic and phenotypic information are analyzed together in one single mixed linear model including all genotyped and non genotyped animals which avoid bias due genomic selection. These approaches are appealing alternatives since they benefit to non genotyped animals, in comparison with the blending approach implemented by Van Raden *et al.* (2009) to compute GEBV from DGV (as described in Section 2-1.3). The approach proposed by

Stoop *et al.* (2011) is particularly interesting as it deals partly with double counting of genomic information for sires with many genotyped sons (issue 2). They are easy to implement without deep modifications of existing evaluation models and software. Several countries have already implemented (or are just about to implement) such multi-step approaches in their national evaluations, including VIT in Germany, MTT in Finland and CRV in the Dutch-Flemish evaluation.

Nevertheless, whatever the method (Ducrocq and Liu, 2009, Liu *et al.*, 2009, Mäntysaari and Strandén, 2010, Stoop *et al.*, 2011, Van Raden *et al.*, 2009), inconsistencies in reliabilities exist and many parameters need to be assumed known (correlations, genomic reliability, β…). None of these methods deal with the complex dependency between input data from genetic and genomic evaluations (issue 3).

- **A method to properly propagate genomic information to the whole population: the single step approach**

With the single step approach suggested by Mistzal *et al.* and Christensen and Lund (2010), the problem of dependency of evaluations (issue 3) or of double-counting of genomic information from relatives (issue 2) disappear.

The principle is that all types of information, i.e., genotypes, pedigree and phenotypes, are considered in a single analysis to simultaneously estimate a single breeding value for each genotyped and non-genotyped animal. The method was described in Section 2-1.3.

The core of the method is the construction of a relationship matrix which combines the classical relationship coefficients with the genomic ones, in such a way that all sources of information are properly distributed and weighed (issue 2 and 3). It especially avoids the use of gEDC.

As all genotyped and phenotyped animals, including the culled candidates, are evaluated at the same time, this approach automatically accounts for genomic selection.

However, the method still requires the knowledge (assumption) of the fraction of the genetic variance explained by the markers (issue 1). Various scalings of the genomic relationship matrix, **G**, were proposed depending on the assumed allele frequencies (Aguilar *et al.*, 2010, Christensen and Lund, 2010, Forni *et al.*, 2011, Vitezica *et al.*, 2011). With improper **G**, genomic information might still be overestimated.

The distribution of genomic information to the whole population including genotyped and non genotyped animals faces three problems. None of the approaches, neither multi-step nor single step, entirely solve the issue of weight given to genomic information compared with the polygenic one (issue 1). But, all of these methods do account for genomic selection! The main advantage of single step approaches over the multi-step ones is that the genomic information can be correctly distributed (issue 2) and the redundancy in the estimation of breeding values can be avoided by the simultaneous use of all types of information (issue 3).

7-3.2. POSSIBLE IMPLEMENTATIONS OF THESE APPROACHES IN ROUTINE EVALUATIONS

- **Computational issues and software adaptations**

The single step approach may face computational problems, in particular to create and invert **G** or **H** if the number of genotyped animals becomes very large. It still needs to be adapted to the wide variety of genetic evaluation models currently used (e.g., threshold models, test-day models, etc…). This approach is computationally costly and requires demanding software adaptations. Legarra and Ducrocq (Ducrocq and Legarra, 2011, 2011) proposed an iterative implementation of the single step approach for genomic evaluation which preserves existing genetic evaluation models and software for classical and genomic prediction.

They start by defining a genetic model equivalent to the single step genetic model. The additive genetic value of an animal (u_i) is broken down into two components, a regular one which corresponds to solutions obtained from regular BLUP evaluations (u_i*) ignoring genomic information and a "strictly" genomic deviation (d_i), which can be derived contrasting classical and genomic solutions. The vector **d** includes genomic contributions for non genotyped ($\mathbf{d_1}$) and genotyped ($\mathbf{d_2}$) animals: $\mathbf{d_1}$ is obtained by regression on $\mathbf{d_2}$ using as regression coefficients a function of blocks of the relationship matrix, **A**:

$$\mathbf{d}_1 = \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{d}_2 \quad [36]$$

where *1* and *2* refer to non genotyped and genotyped animals respectively.

The genetic model can be written:

$$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1\\ \mathbf{X}_2\end{bmatrix}\mathbf{b} + \begin{bmatrix}\mathbf{W}_1 & \mathbf{0}\\ \mathbf{0} & \mathbf{W}_2\end{bmatrix}\begin{bmatrix}\mathbf{u}_1^*\\ \mathbf{u}_2^*\end{bmatrix} + \begin{bmatrix}\mathbf{W}_1\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\\ \mathbf{W}_2\end{bmatrix}\mathbf{d}_2 + \begin{bmatrix}\mathbf{e}_1\\ \mathbf{e}_2\end{bmatrix} \quad [37]$$

The Mixed Model Equations based on this new parameterization (see Ducrocq and Legarra (2011, 2011) for details) can be manipulated to get:

$$\begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 + \mathbf{X}_2'\mathbf{X}_2 & \mathbf{X}_1'\mathbf{W}_1 & \mathbf{X}_2'\mathbf{W}_2 & \mathbf{0} \\ \mathbf{W}_1'\mathbf{X}_1 & \mathbf{W}_1'\mathbf{W}_1 + \alpha_u\mathbf{A}^{11} & \alpha_u\mathbf{A}^{12} & \mathbf{0} \\ \mathbf{W}_2'\mathbf{X}_2 & \alpha_u\mathbf{A}^{21} & \mathbf{W}_2'\mathbf{W}_2 + \alpha_u\mathbf{A}^{22} & -\alpha_u\mathbf{A}_{22}^{-1} \\ \mathbf{0} & \mathbf{0} & -\alpha_u\mathbf{A}_{22}^{-1} & \alpha_u\{\mathbf{A}_{22}^{-1} + (\mathbf{G} - \mathbf{A}_{22})^{-1}\} \end{bmatrix} \begin{bmatrix}\mathbf{b}\\ \mathbf{u}_1\\ \mathbf{u}_2\\ \mathbf{d}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1'\mathbf{y}_1 + \mathbf{X}_2'\mathbf{y}_2\\ \mathbf{W}_1'\mathbf{y}_1\\ \mathbf{W}_2'\mathbf{y}_2\\ \mathbf{0}\end{bmatrix} \quad [38]$$

The equivalence with the single step approach can be demonstrated by absorbing the last equation ($\mathbf{d_2}$) into the other ones. The principle of the iterative approach is to split this system in two to be solved iteratively:

- A regular genetic evaluation system with a modification of the right hand side of genotyped animals:

$$\begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 + \mathbf{X}_2'\mathbf{X}_2 & \mathbf{X}_1'\mathbf{W}_1 & \mathbf{X}_2'\mathbf{W}_2 \\ \mathbf{W}_1'\mathbf{X}_1 & \mathbf{W}_1'\mathbf{W}_1 + \alpha_u\mathbf{A}^{11} & \alpha_u\mathbf{A}^{12} \\ \mathbf{W}_2'\mathbf{X}_2 & \alpha_u\mathbf{A}^{21} & \mathbf{W}_2'\mathbf{W}_2 + \alpha_u\mathbf{A}^{22} \end{bmatrix} \begin{bmatrix}\mathbf{b}\\ \mathbf{u}_1\\ \mathbf{u}_2\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1'\mathbf{y}_1 + \mathbf{X}_2'\mathbf{y}_2\\ \mathbf{W}_1'\mathbf{y}_1\\ \mathbf{W}_2'\mathbf{y}_2\end{bmatrix} + \begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \alpha_u\mathbf{A}_{22}^{-1}\widehat{\mathbf{d}}_2\end{bmatrix} \quad [39]$$

Solutions for the fixed, polygenic ($\boldsymbol{u_1}$) and genomic ($\boldsymbol{u_2}$) effects can be computed from the existing evaluations and software. $\boldsymbol{d_2}$ is assumed to be zero at the first iteration.

- A genomic evaluation or more exactly the estimation of $\mathbf{d_2}$ using a genomic covariance matrix $\mathbf{G}$=var(u) of the HMME: $\alpha_u \mathbf{A}_{22}^{-1} \widehat{\mathbf{d}_2} = \alpha_u (\widehat{t_A} - \widehat{t_G})$ with $\widehat{t_A} = \mathbf{A}_{22}^{-1} \widehat{\mathbf{u}_2}$ and $\widehat{t_G} = \mathbf{G}^{-1} \widehat{\mathbf{u}_2}$.
- t_A and t_G need to be solved to get d_2. Alternatively, $\boldsymbol{u_2}$ used in $\widehat{t_G} = \mathbf{G}^{-1} \widehat{\mathbf{u}_2}$ could be the genomic solutions of a G-BLUP approach or from any genomic evaluation, using DYD derived from the first system.

Such strategy involves moderate modifications of the existing BLUP software used for national genetic evaluations and should be compatible with most genomic evaluation methods.

- **Balance between conceptual issues and ease of implementation: short and long term opportunities**

There are now two alternative strategies to account for genomic selection in routine genetic evaluations.

In theory, a single step approach would provide more accurate genetic evaluations but its direct implementation requires many changes in the genetic evaluation routines which may not be envisioned in short term. However, the need for preventing the creation of bias is urgent: the first recorded daughters of young sires selected after a genomic selection step can be included in genetic evaluations 3 to 4 years after the first implementation of genomic evaluations and selection (Section 5-1.3). The computational strategies proposed by Legarra and Ducrocq (2011, 2011) would provide the advantages of the single step approach and possibly a faster implementation as it only requires relatively reduced software adaptations.

Even though multi-step approaches do not perfectly address the compatibility of genomic information with the classical genetic evaluations, they are relatively easier to implement as they require minor adaptations of the current methods and software. Among multi-step approaches, considering genomic pseudo-performances as a correlated trait to the observed performances under a bivariate model is probably the most appealing approach. To avoid double counting of information, either pedigree information can be removed from the genomic part and/or genomic information can be bounded.

In any case, the chosen strategy has to be tested to check its compatibility with the national data quality requirements, e.g., national certification processes, if any.

- **Perspectives from countries without genomic evaluations**

The single step strategy can only be chosen to account for bias due to genomic selection if the country is implementing genomic evaluations. In case of importing countries without own genomic evaluations, only multi-step approaches can be considered. Conceptually, they could include genomic pseudo-performances after being de-regressed from international breeding

value estimation. In theory, the GMACE method (implemented by Interbull) could provide pertinent GEBV and reliabilities but it assumes that GEBV for all culled and selected candidates were sent to Interbull by the exporting countries. This would be far from reality and complex to implement, particularly if the country is importing sires from various places.

- **Compatibility with Interbull validation tests**

For countries participating in international evaluations, any modification in national evaluation systems should be tested by Interbull. National proofs should first pass the validation tests before being included in MACE (test) runs.

Interbull validation methods I, II or III test the consistency of national genetic evaluations over successive evaluations. The validation methods test whether extra information from additional lactations and/or from new recorded daughters modifies EBV in a systematic way, revealing an incorrect genetic trend. This is checked looking at the regression coefficient of the bulls DYD per year on year of use. These regression coefficients should be close to zero. On the other hand, MACE test runs monitor the consistency of international genetic evaluations over successive evaluations.

If genomic selection is now accounted for in national genetic evaluations, all EBV will include genomic information whatever the method (directly including pseudo-performances of genotyped candidates in BLUP or as a correlated trait, or using a single step approach). The question is whether or not the current Interbull tests are still suitable with this additional genomic information.

- **Missing information to account for a genomic selection step**

The success of the described methods in avoiding bias due to genomic selection depends on the availability of information about culled candidates. All genomic breeding values used for selection should be centralized in the national databases as it was usually done for records to run classical genetic evaluations. However, several barriers may prevent this centralization of all data:

- ✓ If animals are culled very early in life, they may not be registered in the national databases. Their pedigrees may not be known early enough and may not enter any evaluation system afterwards. This is especially relevant for foreign animals or embryos from foreign parents.
- ✓ Genomic evaluations can be performed at a computing center apart from where official national genetic evaluations are done (e.g., in New-Zealand).
- ✓ If genomic breeding values are especially used as a pre-selection tool, one could imagine that AI companies would run their own in-house genomic evaluations. In such a case, where GEBV would be widespread over several centers, it would be difficult to encourage sharing of the information on culled candidates and to be sure that 100% of this information would be collected.
- ✓ Chips with a low density of markers dedicated to the selection of a restricted number of traits could be differently developed across and within countries. This is

appealing as a less expensive alternative to a high density chip. With the development of a low density chip (e.g., the low density chip of Illumina inc. called Infinium BovineLD BeadChip), such dedicated chips are less interesting from a technical point of view but they are still appealing from a marketing point of view (some have begun to develop them, e.g., in beef cattle). Gathering and making this information comparable to the current genomic one would also increase the difficulties with evaluating the entire population with respect to the data on which selection decisions are based.

- ✓ AI companies from different countries can collaborate and exchange genotypes (e.g., AMELIS in France with CRI in the USA). If all the candidates are evaluated in the countries involved (e.g., in France and in the USA), there is no problem. But, if only the genomic breeding values for the selected candidates are transferred (or none of them in the worst case) to the other national computing center(s) for classical evaluations, it becomes an issue. If selection decisions are based on genomic breeding values computed on the foreign scale, it will also be a problem, as the information on the actual selection criteria will be ignored by the national computing center.
- ✓ Countries without genomic evaluations can import foreign bulls based on genomic information. If these genomic breeding values are only available on the foreign scale, the process of selection cannot be considered in their own national genetic evaluations.

- **Information impossible to handle: an approximate approach to account for bias**

If the individual data necessary to account for genomic selection is available, the amount of data will rapidly increase over time. Indeed, more and more males and females are expected to be widely genotyped and the number of genomic breeding values from culled animals may become too large to be handled because of storage or computing requirements. This is particularly likely to occur at the international level. In such a case, a method which summarizes the information and does not require the storage of all the genomic breeding values could be considered.

It was suggested to define new fixed effects such as contemporary groups to account for missing pedigree, for use of highly selected foreign sires or for heterogeneity of residual variance (Section 5-2.1) as used to deal with past situations of bias in genetic evaluations.

The problem of the existence of various subpopulations in classical genetic evaluations was especially addressed for sires imported from different countries. To account for different genetic means for base animals of different origins in national and international evaluations, a proper grouping is now routinely implemented according to sex, country of origin and birth year of animals with unknown parents. A fixed effect of unknown parents groups is then added to the animal model.

Cohorts of genomically selected sires are characterized by a non trivial distribution of MS terms with a mean and variance which can vary from one cohort to another. All contemporary

candidates having undergone genomic selection, if clearly identified, could be grouped by cohorts. Adapting the method of Quaas for unknown parents groups (1988), distinct mean contribution could be computed. They could be estimated from the entire data set of genotyped animals or derived from the true or approximate knowledge of the selection intensity. Assuming a truncation selection based on MS contribution or on EBV, mean and variance of MS terms among selected animals can be approximated. This idea was presented by Patry and Ducrocq (2009) at the Interbull workshop in Uppsala (Sweden) but was neither tested nor implemented. It relies on many assumptions and resulting inaccuracies could accumulate over generations.

This approach has several advantages. Individual data are no longer necessary. Nevertheless, the genetic model should account for genomic selection by defining fixed effects for contemporary groups of genomically selected sires. There are no needs to combine genomic with polygenic information. However, several hypotheses have to be assumed (truncation selection, on which traits, how to compute selection intensity, etc…). This area clearly needs more research work.

7-3.3. TECHNIQUES TO AVOID BIAS PROPAGATION IN INTERNATIONAL GENETIC EVALUATIONS

- **Recommendations to avoid bias propagation**
  - Taking into account genomic selection at the national level first

To be sure to avoid the propagation of bias at the international level and penalization in international comparisons, an obvious recommendation is to account for a genomic selection step in the national genetic evaluations. Several approaches were described previously: they were shown to be efficient and some of them can be implemented in the very near future.

  - Collecting all information on culled and selected genotyped candidates at Interbull

Even if the national proofs collected by Interbull are unbiased after accounting for genomic selection, it was shown (article III) that it is also essential to provide Interbull with all national proofs for culled and selected candidates. This would avoid a “new” selection bias at the international level. However, all bulls, whether culled or selected, have breeding values which are GEBV and include genomic information. Including GEBV into international evaluations raises two questions. The first one is about the further dependence of genomic and classical evaluations as described in this chapter (Section 7-3.1). The second is about including GEBV into international evaluations as described by Van Raden and Sullivan (2010) (Section 6-1.2).

- **Bias detection**

  - Aims of detection tests:

The quality of international genetic evaluations depends on the quality of the data sent to Interbull. It is therefore relevant to try to detect bias due to genomic selection in classical evaluations. The first objective of such detection would be to eliminate biased national evaluations from the international computations and avoid the propagation of bias to other populations. Detecting bias due to genomic selection would also encourage countries to account for it at the national level and help them ensure that optimal evaluation methods are chosen.

  - Four possible sets of national proofs:

National data sets sent by the countries to participate in the international evaluations can be grouped in four different types, corresponding to the following situations:

- Case 1: Countries do not have their own genomic evaluation and do not implement any selection based on genomic information. All data on culled and selected candidates (after progeny testing) are sent ;
- Case 2: Countries have their own genomic evaluation and implement a genomic selection step which is accounted for in the national evaluation. They provide Interbull with all the data on selected and culled candidates. In such a case, they send unbiased and complete national proofs;
- Case 3: Countries have their own genomic evaluation and implement a genomic selection step which is accounted for in the national evaluation. But they provide Interbull with the data on selected bulls only. In such a case, they send unbiased but incomplete national proofs;
- Case 4: Countries have their own genomic evaluation and implement a genomic selection step which is not accounted for in national evaluations. They provide Interbull with the data on selected bulls only. In such a case, they send biased and incomplete national proofs;

In contrast with case 4, no bias need to be detected at the national and international levels in cases 1, 2 and 3.

  - Possible checks to detect bias:

As shown earlier, the magnitude and direction of bias are quite unpredictable which makes difficult the development of a proper test. Only the initial cause of bias is known – the deviation from the expected distribution of MS terms – and therefore checking the characteristics of the MS distribution seems a natural thing to do. The estimated MS distribution should be consistent with the true MS distribution if genetic evaluations are not biased.

The true MS mean can never be known but the prediction error variance (PEV) of the MS deviation, which is the variance of the difference between true and estimated MS values ($PEV(\varphi_i) = \text{var}(\varphi_i - \widehat{\varphi_i})$) can be estimated.

In theory, PEV should be computed from inversion of the coefficient matrix. In fact, Fikse et al (2003) proposed a method to estimate it. PEV would be derived from non-linear equations and based on the approximate reliabilities of animals *i* and of their sire and dam.

PEV can also be computed from simulated true MS terms as derived by Lidauer *et al.* (2007) using repeated sampling and from the estimated MS contribution which is easily computed from real data. This is the simple difference between parent average and individual breeding values based on national proofs.

The various estimations of PEV could be compared to each other and for different subsets of populations (e.g., genotyped versus non genotyped animals). Any inconsistency could help in detecting bias. This issue needs to be more widely studied to develop appropriate tests.

Conclusion: Several approaches were described here to combine all sources of information, i.e., genomic information, pedigree-based and phenotypic information. Multi-step as well as single step approaches avoid the danger of biased genetic evaluations due to genomic selection but some concerns about overestimation or double-counting of genomic information remain. In theory, the single step approach should be preferred. In the short term, a national strategy to combine all types of information in genetic evaluations has to be chosen. Multi-step approaches are already implemented and just need to be adapted to include all selection candidates. The iterative single step approaches could be implemented in the mid term. It would quickly avoid the propagation of bias at the international level. The suitability of Interbull validation tests has to be further examined in this new framework. Any strategy must also be tested to check its compatibility with the national and international data quality requirements.

CHAPTER 8 - General Conclusion

8.1. Impact of genomic selection on genetic evaluations

8-1.1. An established risk for genetic evaluations

Genomic selection is being implemented worldwide. This new strategy for dairy cattle breeding schemes will have major consequences on genetic evaluations. If genetic models are not adapted, the selection process is no longer fully described in the evaluation models which is harmful for the estimation of breeding values.

The implementation of genomic selection clearly violates some of the assumptions of mixed linear models for genetic predictions (Chapter 3). Based on our simulation studies, there are strong evidences that genetic evaluations will become biased (Chapter 4). First, the breeding values of the selected candidates are penalized after genomic selection. It is also clear that such bias is transmitted through the relationship matrix and through the genetic variance-covariance matrix in a multi-trait analysis (Chapter 6) so that bias can widely propagate. Such bias would not only be transferred within a population, i.e., from genotyped to non genotyped animals to males and females, but also between traits and country scales. Genomic selection is expected to occur at each generation so that bias will be regenerated and will accumulate over generations.

The simulation studies were implemented on one trait and over one generation only. In fact, it would be difficult to predict bias magnitude and direction in complex but realistic situations. However, bias is certain and the perspectives for genomic selection would only reinforce this trend. Less expensive chips with a lower marker density is likely to increase the number of genotyped males and females. Sequencing data with multi-breed approaches may increase the accuracy of genomic evaluations. Finally, higher selection intensity is expected to reinforce the bias after genomic selection and to participate to losses in selection efficiency.

Facing such a widespread and uncontrolled problem, there is an obvious need to account for genomic selection in the conception of genetic evaluation models.

8-1.2. An urgent need to avoid such bias

Finding out more robust approaches is also urgent. Between the first selection decisions based on genomic information and the inclusion of first records in classical evaluations, the time interval is short: 3 to 4 years (Chapter 5). Obvious but time-consuming steps are required: sharing awareness of the problem, choosing a method to implement in routine evaluations, testing the method at the national level and eventually at the international one. Nevertheless, there are compulsory to fulfill certification procedures, to guarantee the quality of estimated

breeding values as selection tools and to maintain trust among breeding companies and farmers.

As usual, a compromise between optimal properties of the method and time to implement it has to be found. It is clear that genetic evaluations have to be adapted: the challenge for each national computing centre is now to decide on an efficient but relatively quick and easy way to implement a suitable procedure.

8-1.3. A RELATIVELY EASY APPROACH TO IMPLEMENT AT THE NATIONAL LEVEL

- **A valuable concept: single step evaluations**

Data from all genotyped candidates, i.e., selected as well as culled ones, has to be included in genetic evaluations for optimal predictions. A natural way is to analyze all animals at the same time based on their genomic and phenotypic information (Chapter 2). This is the principle of the single step approach (Christensen and Lund, 2010, Misztal *et al.*, 2009). This method naturally corrects the bias due to genomic selection. The genomic information is also properly propagated through the relationship matrix without redundancy of information (Chapter 7).

- **Implementation in national routine evaluations: the iterative single step method**

Computational strategies have been suggested to implement single step approaches (Aguilar *et al.*, 2010, Legarra *et al.*, 2009, Legarra and Ducrocq, 2011) and make optimal theory becomes routine (Section 7-3.2). To implement the so called iterative single step procedure, only moderate software adaptations are required so that the various classical and genomic evaluation models can be both maintained. Multi-step approaches are already implemented and would require even less time to be adapted but they face the problem of redundant genomic information which incorrectly inflates reliabilities and may lead to suboptimal breeding value estimates (Section 7-3.1).

The implementation of the iterative single step approach is highly recommended: if efforts need to be invested to modify routine evaluations, it is better to directly improve existing and well established methods and software.

However, there is still a need to check whether national breeding values are biased due to the genomic selection step before they are included in international evaluations. Further research works are necessary to provide Interbull with new tests to include in their validation process of national evaluations (Chapter 7-3.3).

8.2. A broader scope of the questions related to bias

Throughout this study, various questions were addressed which remain pertinent on a much more general context: they deal with bias but also with accuracy of genetic evaluations.

8-2.1. INTEGRATION OF GENOMIC INFORMATION IN CLASSICAL EVALUATIONS: OBJECTIVES BEYOND THE BIAS ISSUE

Including all genotyped candidates in classical evaluations accounts for selective genotyping and selective phenotyping in genomic and classical evaluations. More broadly, it deals with combining genomic with phenotypic and pedigree-based information. Multi-step and single step approaches have been addressed: 1) to transfer polygenic information from non genotyped to genotyped animals without phenotypes while accounting for the fraction of variance unexplained by molecular markers; 2) to propagate genomic information from genotyped to non genotyped animals to make the whole population benefit from an additional, early and quite accurate information. In both cases, it aims at improving the accuracy of genetic evaluations.

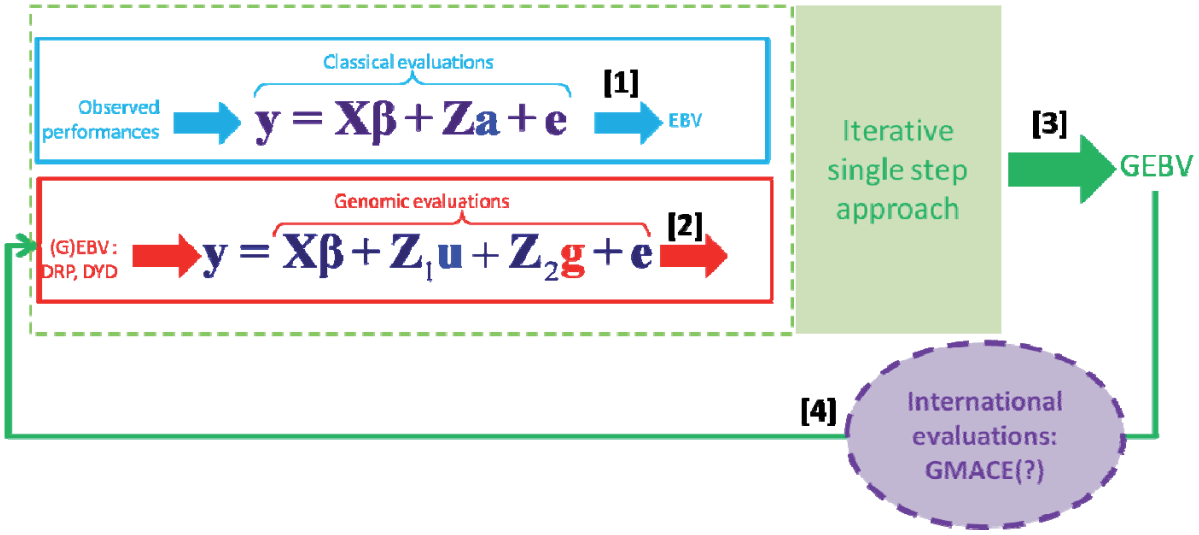
While addressing the problem of bias after a genomic selection step, three general issues came up (Section 7-3.1): 1) the fraction of genetic variance explained by the molecular markers; 2) the double-counting of genomic information in the classical estimation of breeding values; 3) the redundancy of genomic information when the genomic predictions are based on phenotypes already including some genomic information.

If a single step procedure is implemented, the last two issues are avoided. This is not the case with multi-step approaches. Single step approach is definitely an appealing procedure: it allows an optimal use of genomic information while avoiding redundancy.

8-2.2. PERSPECTIVES: GENOMIC INFORMATION PROPAGATED IN ALL ESTIMATIONS OF BREEDING VALUES

Whether single or multi-step approaches are used, all estimated breeding values will necessarily contain genomic information in a near future: GEBV will be delivered for all animals. If single step procedures are chosen, the whole process of genetic evaluation would be modified as described in Figure 2. On the one hand, classical evaluations should be performed (step 1) to deliver EBV based on actual records (i.e., observed performances) as phenotypes. On the other hand, genomic evaluations should be run (step 2) to deliver direct genomic value (DGV) for genotyped animals. The latter are based on phenotypes (DRP or DYD) derived from EBV. Through an iterative single step procedure (step 3), both evaluations converge to deliver GEBV. Consequently, only GEBV can be sent to Interbull for international evaluation and only phenotypes including genomic information (i.e., national or international GEBV) can be used for the future genomic evaluations.

Figure 2: Organization of a genetic evaluation process combining classical, genomic and international evaluations.



Legend: a are additive genetic values; u are breeding values from polygenic information only ; g are breeding values from genomic information only ; EBV for estimated breeding value ; DGV for direct genomic value ; GEBV for genomically enhanced breeing value ; GMACE for genomic MACE.

8-2.3. UNSOLVED ISSUES AT THE INTERNATIONAL LEVEL: INCOMPATIBILITIES WITH MACE

The propagation of genomic information and the fact that only GEBV may be available at Interbull in the near future leads to certain problems.

Countries participating to international evaluations and implementing a genomic evaluation can only deliver to Interbull breeding values with genomic information. However, the regular MACE method is not adapted to GEBV, generating non zero residual variances for sires genotyped in several countries. MACE in its present form can no longer exist for breeds with molecular information. Extension of the single step approach to the international case could be envisioned. Indeed, it is definitely necessary to develop a proper method to perform international evaluations based on GEBV: international GEBV are highly required.

First, there is a need to **compare GEBV on each country scale** to maintain international semen trade based on fair selection tools such as estimated breeding values.

Second, for the creation or the increase of multinational RP, not only genotypes but also phenotypes need to be shared. They are currently in the form of MACE EBV expressed in the country scale of interest. With the wide propagation of genomic information, **GEBV also need to be converted on the relevant country scales** to be shared.

Ideally, Interbull could first manage an international database of genotypes. Depending on political agreements, these genotypes would be available and participating countries could compute GEBV in-house on their own country scale by implementing their own genomic predictions. These GEBV would be mainly useful for multinational RP. These international

GEBV would also provide a tool for international comparisons but in a less extensive way as before: all GEBV could be available on the scales of countries implementing genomic evaluations but not on all the scales. Hence, it would encourage all countries to develop their own genomic predictions.

This clearly shows some limits of genetic evaluations at both the national and the international level: 1) It will be difficult for countries to maintain their classical evaluations without adapting it to genomic information. 2) MACE can no longer exit without adaptations: methods for GEBV conversion on various country scales and tools for international comparison are obviously needed. New and radical developments are expected, and especially at the international level, but depend on the political framework, i.e., the future organization of the world dairy cattle breeding in a changing environment.

8.3. Perspectives for international collaborations in the dairy cattle breeding organization

8-3.1. Economical and technological context: a high competitiveness

In a context of world market liberalization, many changes have been observed in the world dairy cattle breeding organization. Because of an increased competitiveness within and between countries, heavy restructuring has occurred and is still occurring: for example, local breeding cooperatives tend to merge and create larger AI companies.

A technological innovation such as the access to genomic information accelerates these heavy changes: genomic tools have encouraged breeding companies to not only reconsider their breeding design but also the nature of their international relationships.

8-3.2. Opportunities and fears generated by genomic selection

Many changes in methods, practices and strategies are expected in a very short time in order to fully benefit from the exciting opportunities of genomic selection. Genetic gains should be increased but selection on new traits is also expected. Farmers are waiting for new herd management tools. Genomic selection is also seen as a great opportunity for countries without (satisfying) breeding programs, e.g., under harsh environments, to develop one: such promising tools should motivate to extensively record and analyze data

Including genomic information in methods and genomic selection in breeding schemes are also associated with risks, e.g., a risk of lower quality of the estimation of breeding values. Hence, it generates fears: loss of selection efficiency, loss of trust from farmers, financial loss due to higher risk investments. The opportunities for changes are huge but uncertainty increases. Actually, fundamental questions are again raised:

- What is a competitive advantage today: the delay to deliver accurate breeding values, the design of breeding schemes using them, the semen marketing? How can actors be different from the others?
- Which asset will be the most valuable: genotypes, phenotypes, statistical methods, breeding strategies, animals, breeds?
- On which arguments selection decisions will be based? What is the balance between marketing through "brand loyalty" and/or through transparent technical arguments, e.g., breeding values?
- What is the role of livestock breeders in the value chain? Will they actively participate to create genetic progress or will they be simply clients or consumers?

8-3.3. THE ON-GOING CHANGES IN THE ORGANIZATION OF THE WORLD DAIRY CATTLE BREEDING

In this uncertain environment, the global dairy cattle organization is radically changing. Before the genomic turning point, there was one type of breeding values, mostly estimated using one methodology and it was possible to perform fair international comparisons based on international evaluations. The system was clearly structured, in particular, centralized (i.e., with Interbull in the center) and transparent. Today, there are two consortia including the largest Holstein populations and splitting the dairy cattle world into two parts. Interbull is no longer central, its role is no longer so well defined. Many actors still need to find a place in this new structure. The organization of the world dairy cattle breeding has changed but is not yet settled.

8-3.4. NEW MARKET STRATEGIES AND PLAYERS

Such a radical innovation also tends to attract new entrants, e.g., new actors from food or veterinary fields, on the market. However, rivalry among existing competitors is already high.

The possible implementation of genomic selection has increased the differentiation between companies at two levels. On the one hand, there are countries implementing genomic selection versus countries not implementing it. On the other hand, among countries using genomic tools, various breeding designs are possible. In such a case, each company may avoid interactions with the other players to develop a unique and original strategy on the market.

However, the need for accurate genomic predictions has also increased interdependency between international actors. It is crucial to share genotypes and phenotypes and it encourages sharing methods, software and know-how. Collaborations for joint development can get access to more innovative and more efficient selection tools. Existing collaborations and shared history through Interbull have shown the benefits of such a wide cooperation. Moreover, farmers are used to this international network they trust and this is a first barrier to new entrants.

In such a renewed co-opetition framework, international collaborations are essential but it is uncertain whether the future organization of the world dairy cattle breeding will include Interbull in the same way as before.

To conclude, the availability of genomic information has led to a series of changes in the dairy cattle breeding world. This doctoral study has been performed during an exceptional transition period. Genomic selection offers great opportunities but brings a lot of questions to adapt the well established classical system to the use of new genomic information. One key problem among others was the emergence of a bias in genetic evaluations after genomic selection. In fact, it is included in a larger question, i.e., how to combine genomic and polygenic evaluations for the whole population considered. Many approaches are currently under development. One major unsolved difficulty remains the adaptation of methods for international evaluations, even crucial. These are great challenges which need intensive research work. There are also major political issues because of opportunities for new balances between countries and a new world organization of the dairy cattle breeding. In such a context, stakes are high and the future might bring great changes to keep creating genetic gains in a sustainable way.

REFERENCES

1. Aguilar, I., I. Misztal, D. L. Johnson, A. Legarra, S. Tsuruta, and T. J. Lawlor. 2010. A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J Dairy Sci 93:743-752.

2. Boichard, D., B. Bonaiti, A. Barbat, and S. Mattalia. 1995. Three Methods to Validate the Estimation of Genetic Trend for Dairy Cattle. Journal of dairy science 78(2):431-437.

3. Boichard, D., F. Guillaume, A. Baur, P. Croiseau, M. N. Rossignol, M. Y. Boscher, T. Druet, L. Genestout, A. Eggen, L. Journaux, V. Ducrocq, and S. Fritz. 2010. Genomic prediction in French dairy cattle. in Proc. 9th WCGALP, Leipzig, Germany.

4. Bonaiti, B. and D. Boichard. 1995. Accounting for foreign information in genetic evaluation. Interbull Bulletin 11.

5. Bouquet, A., M. J. Sillanpää, and J. Juga. 2011. Estimating inbreeding using dense panels of biallelic markers and pedigree information. in Proc. 62th meeting of EAAP, Stavanger, Norway.

6. Bulmer, M. G. 1971. The Effect of Selection on Genetic Variability. The American Naturalist 105(943):201-211.

7. Christensen, O. and M. Lund. 2010. Genomic prediction when some animals are not genotyped. Genet Sel Evol 42:2.

8. Colleau, J. J. 1989. Impact of the use of bovine somatotropin (BST) on dairy cattle selection. Genet. Sel. Evol. 21(4):479-491.

9. Colleau, J. J., S. Fritz, F. Guillaume, A. Baur, D. Dupassieux, M. Y. Boscher, L. Journaux, A. Eggen, and D. Boichard. 2009. Simulation des potentialités de la sélection génomique chez les bovins laitiers. in 17ème journées 3R. Paris, France.

10. Colombani, C., P. Croiseau, S. Fritz, F. Guillaume, A. Legarra, V. Ducrocq, and C. Robert-Granié. 2011. A comparison of PLS and sparse PLS regressions in genomic selection in French dairy cattle. Page Submitted. J Dairy Sci.

11. Croiseau, P., A. Legarra, F. Guillaume, S. Fritz, A. Baur, C. Colombani, C. Robert-Granié, D. Boichard, and V. Ducrocq. 2011. Fine tuning genomic evaluations in dairy cattle through SNP pre-selection with Elastic-Net algorithm. Genetic research.

12. Daetwyler, H. D., R. Pong-Wong, B. Villanueva, and J. A. Woolliams. 2010. The Impact of Genetic Architecture on Genome-Wide Evaluation Methods. Genetics 185(3):1021-1031.

13. Daetwyler, H. D., B. Villanueva, P. Bijma, and J. A. Woolliams. 2007. Inbreeding in genome-wide selection. J Anim Breed Genet 124:369-376.

14. de Roos, A. P. W. 2011. Genomic selection in dairy cattle. Page 184 in Animal breeding and genetics. Vol. doctor. Wageningen.

15. Dekkers, J. C. M. 2004. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. J Anim Sci 82:313-328.

16. Druet, T., S. Fritz, M. Boussaha, S. Ben-Jemaa, F. Guillaume, D. Derbala, D. Zelenika, D. Lechner, C. Charon, D. Boichard, I. G. Gut, A. Eggen, and M. Gautier. 2008. Fine Mapping of Quantitative Trait Loci Affecting Female Fertility in Dairy Cattle on BTA03 Using a Dense Single-Nucleotide Polymorphism Map. Genetics 178:2227-2235.

17. Ducrocq, V. 2010. Sustainable dairy cattle: illusion or reality? in Proc. 9th WCGALP, Leipzig, Germany.

18. Ducrocq, V., I. Delaunay, D. Boichard, and S. Mattalia. 2003. A general approach for international genetic evaluations robust to inconsistencies of genetic trends in national evaluations. Interbull bulletin 30:101-111.

19. Ducrocq, V., M. Goddard, L. Schaeffer, and G. Banos. 2008. Interbull Scientific Advisory Committee annual report (2007-2008) to the Interbull Steering Committee. Pages 1-4 in Interbull meeting. Niagara Falls, USA.

20. Ducrocq, V. and A. Legarra. 2011. An iterative implementation of the single step approach for genomic evaluation which preserves existing genetic evaluation models and software. in Proc. The 2011 Interbull meeting, Stavanger, Norway.

21. Ducrocq, V. and Z. Liu. 2009. Combining genomic and classical information in national BLUP evaluations. Interbull Bulletin 40:172-177.

22. Ducrocq, V. and E. Santus. 2011. Moving away from progeny testing schemes: consequences on conventional (inter)national evaluations. in Interbull technical workshop. Guelph, Ontario, Canada.

23. Ehsani, A., L. Janss, and O. F. Christensen. 2011. Exploring trend of bias to prediction of genomic breeding values using Bayesian variable selection in the population undergoing selection. in 62th meeting of the EAAP. Stavanger, Norway.

24. Ehsani, A. R., L. Janss, and O. F. Christensen. 2010. Effect of selective genotyping on genomic predictions. in Proc. 9th WCGALP, Leipzig, Germany.

25. Fikse, W. and G. Banos. 2001. Weighting factors of sire daughter information in international genetic evaluations. J Dairy Sci 84:1759 - 1767.

26. Fikse, W. F., L. Klei, Z. Liu, and P. G. Sullivan. 2003. Procedure for validation of trends in genetic variances. Interbull bulletin 31:30-36.

27. Forni, S., I. Aguilar, and I. Misztal. 2011. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. Gen Sel Evol 43:1-7.

28. Foulley, J. L., D. Gianola, M. San Cristobal, and S. Im. 1990. A Method for Assessing Extent and Sources of Heterogeneity of Residual Variances in Mixed Linear Models. J Dairy Sci 73:1612-1624.

29. Fritz, S. and F. Guillaume. 2008. Travaux de cartographie de QTL et développement de la SAM chez les bovins laitiers. "Séminaire génomique", AgroParisTech.

30. Fritz, S., F. Guillaume, P. Croiseau, A. Baur, C. Hoze, R. Dassonneville, M. Y. Boscher, L. Journaux, D. Boichard, and V. Ducrocq. 2010. Implementing genomic selection in the three main French dairy cattle breeds. in Proc. 18ème journées 3R, Paris, France.

31. Gianola, D., R. L. Fernando, and A. Stella. 2006. Genomic-Assisted Prediction of Genetic Value With Semiparametric Procedures. Genetics 173:1761-1776.

32. Goddard, M. and B. Hayes. 2007. Genomic selection. J. Anim. Breed. Genet. 124(6):323-330.

33. Guillaume, F., S. Fritz, D. Boichard, and T. Druet. 2008. Short Communication: Correlations of Marker-Assisted Breeding Values with Progeny-Test Breeding Values for Eight Hundred Ninety-Nine French Holstein Bulls. J Dairy Sci 91:2520-2522.

34. Harris, B. and D. Johnson. 1998. Approximate Reliability of Genetic Evaluations Under an Animal Model. J Dairy Sci 81(10):2723-2728.

35. 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(2):433-443.

36. Henderson, C. R. 1963. Selection index and expected genetic advance. Pages 141-163 in Nat. Acad. Sci., Nat. Res. Council. Vol. 982.

37. Henderson, C. R. 1973. Sire evaluation and genetic trends. J Anim Sci 1973:10-41.

38. Henderson, C. R. 1975. Best Linear Unbiased Estimation and Prediction under a selection model. Biometrics 31(2):25.

39. Henderson, C. R. and R. L. Quaas. 1976. Multiple Trait Evaluation Using Relatives' Records. J Anim Sci 43(6):1188-1197.

40. Hickey, J., M. Keane, D. Kenny, A. Cromie, H. Mulder, and R. Veerkamp. 2008. Estimation of accuracy and bias in genetic evaluations with genetic groups using sampling. J Anim Sci 86(5):1047-1056.

41. Interbull. 1992. Sire evaluation procedures for dairy production traits practised in various countries. Interbull Bulletin 5.

42. Interbull. 2011. Establishing the framework for international genomic evaluations - executive summary.

43. Jairath, L., J. C. Dekkers, L. R. Schaeffer, Z. Liu, E. B. Burnside, and B. Kolstad. 1998. Genetic evaluation for herd life in Canada. J Dairy Sci 81(2):550-562.

44. Jakobsen, J. and F. Fikse. 2009. Sire-Dam pedigree in MACE - Results from a full-scale pilot study. Interbull Bulletin 39:91-98.

45. Kennedy, B. W., L. R. Schaeffer, and D. A. Sorensen. 1988. Genetic Properties of Animal Models. J. Dairy Sci. 71:17-26.

46. Kuhn, M. T., P. J. Boettcher, and A. E. Freeman. 1994. Potential Biases in Predicted Transmitting Abilities of Females from Preferential Treatment. J. Dairy Sci. 77:2428-2437.

47. Lande, R. and R. Thompson. 1990. Efficiency of Marker-Assisted Selection in the Improvement of Quantitative Traits. Genetics 124(3):743-756.

48. Le Bourhis, D. M., E. and P. C. Humblot, W. Ponsart, C. 2010. Bovine embryo genotyping using a 50k single nucleotide polymorphism chip. Reprod. Fertil. Dev. 23(1):197-197.

49. Legarra, A., I. Aguilar, and I. Misztal. 2009. A relationship matrix including full pedigree and genomic information. J Dairy Sci 92:4656-4663.

50. Legarra, A. and V. Ducrocq. 2011. Computational strategies for national integration of phenotypi, genomic and pedigree data in a single-step BLUP. J Dairy Sci:Submitted.

51. Legarra, A., C. Robert-Granié, P. Croiseau, F. Guillaume, and S. Fritz. 2011. Improved Lasso for genomic selection. Genetics Research 93(01):77-87.

52. Lidauer, M., K. Vuori, I. Stranden, and E. A. Mäntysaari. 2007. Experiences with Interbull IV: Estimation of genetic variance. Interbull Bulletin 37:69-72.

53. Liu, Z., F. Seefried, F. Reinhardt, and R. Reents. 2009. A simple method for correting the bias caused by genomic pre-seletion in conventional genetic evaluation. Interbull Bulletin 40:184-188.

54. Lund, M. S., A. P. W. de Roos, A. G. de Vries, T. Druet, V. Ducrocq, S. Fritz, F. Guillaume, B. Guldbrandtsen, L. Liu, R. Reents, C. Schrooten, M. Seefried, and G. Su. 2010. Improving genomic predictions by Eurogenomics collaborations. in Proc. 9th WCGALP, Leipzig, Germany.

55. Lund, M. S., A. P. W. de Roos, A. G. de Vries, T. Druet, V. Ducrocq, S. Fritz, F. Guillaume, B. Guldbrandtsen, Z. Liu, R. Reents, C. Schrooten, M. Seefried, and G. Su. 2011. Common reference of four European Holstein populations increases reliability of Genomic Predictions. Genet Sel Evol.

56. Mackinnon, M. J. and M. Georges. 1992. The Effects of Selection on Linkage Analysis for Quantitative Traits. Genetics 132(4):1177-1185.

57. Maher, B. 2008. Personal genomes: The case of missing heritability. Nature 546:18-21.

58. Mallinckrodt, C., B. Golden, and R. Bourdon. 1995. The effect of selective reporting on estimates of weaning weight parameters in beef cattle. J Anim Sci 73(5):1264-1270.

59. Meuwissen, T. and M. Goddard. 2001. Prediction of identity by descent probabilities from marker-haplotypes. Genet Sel Evol 33:605 - 634.

60. 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(4):1819-1829.

61. Misztal, I., A. Legarra, and I. Aguilar. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. J Dairy Sci 92:4648-4655.

62. Mrode, R. 2005. Linear models for the prediction of animal breeding values. Second edition ed. CABI publishing, Wallingford, UK.

63. Mäntysaari, E., Z. Liu, and P. VanRaden. 2010. Interbull validation test for genomic evaluations. Interbull bulletin 41:17-22.

64. Mäntysaari, E. A. and I. Strandén. 2010. Use of bivariate EBV-DGV model to combine genomic and conventional breeding value estimations. in Proc. 9th WCGALP, Leipzig, Germany.

65. Nilforooshan, M. A. 2011. Multiple-trait multiple country genetic evaluation of fertility traits in dairy cattle. in Veterinary medecine and animal science. Vol. phd. Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden.

66. Nilforooshan, M. A., B. Zumbach, J. Jakobsen, A. Loberg, H. Jorjani, and J. Dürr. 2010. Validation of National Genomic Evaluations. Interbull Bulletin 42:56-61.

67. Patry, C. and V. Ducrocq. 2009. Bias due to genomic selection. Interbull Bulletin 39:77-82.

68. Patry, C. and V. Ducrocq. 2011a. Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. Genet Sel Evol 43:30.

69. Patry, C. and V. Ducrocq. 2011b. Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. J Dairy Sci 94(2):1011-1020.

70. Patry, C., H. Jorjani, and V. Ducrcoq. Implementation of genomic selection at national level: impact of pre-selected and biased national BLUP evaluations on international genetic evaluations. J Dairy Sci.

71. Pedersen, G. A., J. Pedersen, U. S. Nielsen, and P. Madsen. 1999. Experiences of blending foreign information in the national genetic evaluations. Interbull Bulletin 22:61-65.

72. Pollak, E. J. and R. L. Quaas. 1981. Monte Carlo Study of Genetic Evaluations Using Sequentially Selected Records. J Anim Sci 52:257-264.

73. Pollak, E. J., J. van der Werf, and R. L. Quaas. 1984. Selection bias and multiple trait evaluation. J Dairy Sci 67:1590-1595.

74. Powell, R. L. and H. D. Norman. 2006. Major Advances in Genetic Evaluation Techniques. J Dairy Sci 89:1337-1348.

75. Quaas, R. L. 1988. Additive Genetic Model with Groups and Relationships. J Dairy Sci 71:1338-1345.

76. Reverter, A., B. Golden, R. Bourdon, and J. Brinks. 1994. Technical note: detection of bias in genetic predictions. J Anim Sci 72(1):34-37.

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

78. Schaeffer, L. R. 2006. Strategy for applying genome-wide selection in dairy cattle. J Anim Breed Genet 123(4):218-223.

79. Schaeffer, L. R., F. S. Schenkel, and L. A. Fries. 1998. Selection Bias on Animal Model Evaluation. in Proc. 6th WCGALP, Armidale, Australia.

80. Searle, S. 1982. Matrix Algebra Useful for Statistics (Wiley Series in Probability and Statistics). Wiley-Interscience.

81. Sigurdsson, A. and G. Banos. 1995. Dependent Variables in International Sire Evaluations. Acta Agriculturae Scandinavica, Section A - Animal Science 45(4):209-217.

82. Simianer, H., J. Chen, and M. Erbe. 2011a. Animal breeding in the genomis era: challenges and opportunities for the maintenance of genetic diversity. in 62th meeting of EAAP. Stavanger, Norway.

83. Simianer, H., M. Erbe, S. König, and E. C. G. Pimentel. 2011b. Estimation of the effective number of genes underlying quantitative traits based on chromosomal partitioning of the genomic variance. in Proc. 62thannual meeting of EAAP, Stavanger, Norway.

84. Smith, C. 1967. Improvement of metric traits through specific. Anim. Prod. 9:349-358.

85. Soller, M. and S. Beckmann. 1983. Genetic polymorphism in variety identification and genetic improvement. Theor. Appl. Genet. 67:25-33.

86. Sorensen, D. A. and B. W. Kennedy. 1983. The use of the relationship matrix to account for genetic drift variance in the analysis of genetic experiments. Theor. Appl. Genet. 66:217-220.

87. Sorensen, D. A. and B. W. Kennedy. 1984. Estimation of Response to Selection Using Least-Squares and Mixed Model Methodology. J. Anim Sci. 58:1097-1106.

88. Stoop, W. M., H. Eding, M. L. van Pelt, and G. de Jong. 2011. Combining genomic and conventional data in the Dutch national evaluation. Page http://www.interbull.org/images/stories/Stoop.pdf in Proc. Interbull meeting, Stavanger, Norway.

89. Thompson, R. and K. Meyer. 1986. A review of theoretical aspects in the estimation of breeding values for multi-trait selection. Livest Prod Sci 15:299-313.

90. Tyriserva, A.-M., M. H. Lidauer, F. Fikse, and E. A. Mäntysaari. 2011. Simulation study on Mendelian sampling variance tests. in Interbull meeting. Stavanger, Norway.

91. Van der Beck, S. 2007. Effect of Genomic Selection on National and International Genetic Evaluations. Interbull Bulletin 37:115-118.

92. Van Raden, P. 2008. Efficient methods to compute genomic predictions. J Dairy Sci 91:4414 - 4423.

93. Van Raden, P. and P. Sullivan. 2010. International genomic evaluation methods for dairy cattle. Genet Sel Evol 42:7.

94. Van Raden, P., C. Van Tassell, G. Wiggans, T. Sonstegard, R. Schnabel, J. Taylor, and F. Schenkel. 2009. Invited review: reliability of genomic predictions for North American Holstein bulls. J Dairy Sci 92:16-24.

95. Van Raden, P. M. and G. R. Wiggans. 1991. Derivation, Calculation, and Use of National Animal Model Information. J Dairy Sci 74:2737-2746.

96. Van Vleck, L. D. 1987. Contemporary Groups for Genetic Evaluations. J Dairy Sci 70(11):2456-2464.

97. Vinson, W. E. 1987. Potential Bias in Genetic Evaluations from Differences in Variation Within Herds. J Dairy Sci 70(11):2450-2455.

98. Vitezica, Z. G., I. Aguilar, and A. Legarra. 2010. One-Step vs. Multi-Step Methods For Genomic Prediction in Presence Of Selection. in 9th World Congress of Genetics Applied to Livestock Production. Leipzig, Germany.

99. Vitezica, Z. G., I. Aguilar, I. Misztal, and A. Legarra. 2011. Bias in genomic predictions for populations under selection. Genet Res:1-10.

100. Weigel, K. A. and G. Banos. 1997. Effect of Time Period of Data Used in International Dairy Sire Evaluations. J Dairy Sci 80:3425-3430.

Table of illustrations

Tables

Table 1: Measures of error and precision in the estimation of breeding values..... 37

Table 2: Accounting for selection bias or other sources of bias in BLUP evaluations..... 63

Figures

Figure 1: Propagation of bias throughout the genetic evaluation process after a selection step (2) based on genomic information (1)..... 113

Figure 2: Organization of a genetic evaluation process combining classical, genomic and international evaluations. 131

Publications

Peer-reviewed publications

Patry, C. and V. Ducrocq, 2011.
Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle
J Dairy Sci 94:1011-1020

Patry, C. and V. Ducrocq, 2011.
Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle
Genet Sel Evol 43:30

Patry, C., H. Jorjani and V. Ducrocq.
Implementation of genomic selection at national level: impact of pre-selected and biased national BLUP evaluations on international genetic evaluations
Submitted (J Dairy Sci) on September, the 27th, 2011

Other publications

Patry, C. and V. Ducrocq, 2009.
Bias due to genomic selection
Interbull Bulletin 39

Patry, C. and V. Ducrocq, 2009.
Evidence of a bias in genetic evaluation due to genomic selection
Interbull Bulletin 40

Patry, C. and V. Ducrocq, 2010.
An approach to account for selection bias in national evaluations due to genomic selection
Proc. 9th WCGALP, Leipzig, Germany

Patry, C., H. Jorjani and V. Ducrocq, 2011.
Implementation of genomic selection at national level: impact of pre-selection and biased national BLUP evaluations on international genetic evaluations
Interbull Meeting, Stavanger, Norway

Co-authored publications

Guillaume, F., Fritz, S., Croiseau, P., Legarra, A., Robert-Granié, C., Colombani, C., Patry, C., Boichard, D., Ducrocq, V., 2009

Modèles d'évaluation génomique: application aux populations bovines laitières françaises
16ème journées 3R, pp. 399-406, Paris , France

Croiseau, P., Colombani, C., Legarra, A., Guillaume, F., Fritz, S., Baur, A., Dassonneville, R., Patry, C., Robert-Granié, C., Ducrocq, V.

Improving genomic evaluation strategies in dairy cattle through SNP pre-selection
Proc. 9th WCGALP., Leipzig, Germany

Résumé

C'est en 2008 que les premières évaluations génomiques ont été mises en place chez les bovins laitiers. Elles reposent sur l'exploitation de l'information dense des marqueurs moléculaires sur le génome. En 2011, au moins 16 pays et 7 races disposent de ces nouveaux index pour une sélection précoce des jeunes taureaux. Les perspectives sont grandes pour accélérer le progrès génétique sur un plus grand nombre de caractères. En effet, on dispose maintenant d'une grande précision des index dès la naissance, pour les mâles comme pour les femelles, et même pour des caractères faiblement héritables. Les décisions de sélection peuvent dès lors reposer sur ces nouveaux outils intégrant une information génomique et cette étape, dite de sélection génomique, peut remplacer le long processus de testage sur descendance. Par conséquent, l'organisation des schémas de sélection est en train de se transformer radicalement.

L'estimation des valeurs génétiques repose sur des modèles statistiques particuliers, que sont les modèles linéaires mixtes. Classiquement, ces modèles analysent une information phénotypique. Les évaluations génétiques utilisent le plus souvent la méthode du BLUP (meilleure prédiction linéaire non biaisée) à l'échelle nationale ou son extension à l'échelle internationale, soit la méthode du MACE (Multiple Across Country Evaluation). Leurs propriétés sont optimales quand l'ensemble des informations ayant servi à la sélection est considéré. Dans les équations du BLUP par exemple, les performances de tous les animaux candidats à la sélection doivent êtres inclues. Si la sélection repose sur des index utilisant l'information génomique et non sur des index classiques calculés après testage sur descendance, seuls les taureaux sélectionnés auront des filles avec performances. Dans les modèles d'évaluation classique, on sait que des informations manquantes peuvent conduire à des estimations génétiques biaisées. On appelle biais toute surestimation ou sous-estimation systématique des valeurs prédites.

Les évaluations génétiques classiques fournissent des phénotypes corrigés qui sont utilisés en évaluation génomique ce qui rend leur calcul indispensable et stratégique. Etant donnée la rapide ampleur de la sélection génomique chez les bovins laitiers, il est important de considérer le risque de biais pour maintenir des évaluations classiques de qualité.

Le premier objectif de cette thèse était de mesurer ce biais dans les évaluations génétiques classiques. Les effets de la sélection génomique sur les données servant aux évaluations classiques ont été simulés. Pour cela, des données réelles issues de la population Holstein française ont été utilisées. Deux situations ont été comparées, celle où les informations sont complètes, c'est-à-dire après un programme de testage sur descendance et celle où les informations sont incomplètes après une étape de sélection génomique. Des valeurs génétiques vraies ainsi que des index « génomiques » ont été simulés pour l'ensemble des jeunes taureaux identifiés comme candidats à la sélection. Des valeurs génétiques vraies ainsi que des performances ont aussi été simulées pour caractériser les filles des taureaux sélectionnés. Ces deux scénarios et leurs évaluations génétiques ont été répétés pour pouvoir calculer un biais moyen égal à la moyenne des différences entre les valeurs génétiques

estimées et les valeurs génétiques vraies. Des indicateurs de précision des évaluations ont aussi été comparés. Différentes héritabilités des caractères, différentes précisions des évaluations génomiques et proportions d'individus retenus, soit différentes intensités de sélection ont été testées. Les simulations ont montré que les taureaux retenus après une étape de sélection génomique étaient fortement impactés : les solutions du BLUP sous-estimaient les valeurs génétiques vraies. Les prédictions étaient aussi moins précises. Les simulations ont montré que leurs filles avaient aussi des estimations biaisées, le biais étant transmis par la matrice de parenté du BLUP. On peut finalement s'attendre à ce que les classements des jeunes taureaux ne soient pas optimaux.

Il est donc indispensable de prendre en compte l'étape de sélection génomique dans les modèles d'évaluations classiques. Les candidats éliminés n'ont pas de performances réelles mais on dispose d'une information génomique pour les caractériser. On propose de dé-régresser les index génomiques pour les considérer comme des pseudo-performances à introduire dans les évaluations classiques. Elles nécessitent cependant d'être pondérées par un terme qui dépend de la précision des évaluations génomiques et qui traduit la quantité d'information d'origine génomique. Cette méthode a été testée par simulation. La même structure de données utilisée pour mesurer le biais a été de nouveau utilisée. Aux deux premiers cas de figure (des performances pour tous les candidats ou pour les animaux sélectionnés seulement), deux nouveaux scénarios ont été simulés : des pseudo-performances génomiques pour l'ensemble des candidats ou seulement pour les candidats retenus. La méthode s'est avérée très satisfaisante dans le seul cas où tous les jeunes taureaux, les sélectionnés comme les éliminés, étaient inclus dans l'évaluation classique avec des pseudo-performances génomiques pondérées. Non seulement le biais est alors inexistant mais l'estimation génétique des jeunes taureaux gagne en précision : outre l'information phénotypique classique, l'information génomique contribue dès lors au processus d'évaluation et à sa précision.

Cette méthode ne peut cependant pas être mise en place directement dans les évaluations classiques de routine et nécessite des adaptations. L'information génomique dans les évaluations classiques est transmise à tous les animaux apparentés par la matrice de parenté. Or, quand un taureau est génotypé ainsi que un ou plusieurs de ses descendants, l'information qui circule est redondante. En effet, la connaissance du génotype d'un fils n'apporte à priori aucune information génomique supplémentaire dès lors que le père est génotypé et vice versa. La précision des index classiques est alors surestimée. Le calcul des poids associés aux pseudo-performances génomiques est donc à adapter. Plusieurs approches ont été comparées pour mieux définir l'importance relative de la contribution génomique par rapport à l'information classique des performances et du pedigree. Certaines reposent sur le calcul de coefficients de détermination plus réalistes à partir de la précision des index génomiques, supposée connue. D'autres proposent d'analyser simultanément les index ou performances classiques avec les index ou pseudo-performances génomiques dans un modèle bi-caractère. Dans tous les cas, l'information génomique bénéficie à tous les animaux de la population, même non génotypés.

Quelle que soit la méthode de prise en compte de l'information génomique dans les évaluations classiques, tous les index obtenus contiennent une information génomique. Or ces index (sous forme de-régressée ou utilisés pour le calcul de performances moyennes corrigées des filles) alimentent la population de référence pour les futures prédictions génomiques. La même information est donc présente à la fois dans les observations comme dans les variables à prédire. Une telle dépendance entre évaluations peut nuire à la qualité des futures prédictions génomiques.

Une évaluation génomique en une seule étape ne présente pas cet inconvénient. Elle intègre tous les animaux de la population à évaluer. Les animaux génotypés ou non sont analysés simultanément. L'information génomique est automatiquement pondérée et propagée à l'ensemble de la population. Les évaluations génétiques sont alors toutes plus précises et non impactées par une étape de sélection génomique. Mais cette méthode pourrait s'avérer compliquée à mettre en œuvre pour des modèles sophistiqués ou pour de très grand nombre d'animaux génotypés. Elle nécessiterait de nombreuses adaptations des programmes d'évaluation en place et donc du temps. Des adaptations calculatoires de cette méthode ont cependant été proposées. Leur principe repose sur une résolution par itérations successives entre les équations des modèles d'évaluation classique d'une part et celles de l'évaluation génomique déjà en place d'autre part. Ce principe permettrait de répartir correctement les différentes sources d'information entre tous les individus de la population et n'exigerait que des adaptations relativement modérées des programmes d'évaluation existants. Ce procédé doit être testé sur données réelles mais pourrait être mis en place sans délais excessifs.

Il est en effet essentiel de trouver des solutions rapidement. Entre la première sélection sur information génomique et les premiers enregistrements de performances prises en compte dans les évaluations classiques, le délai n'est que de 3 à 4 années. En France, la sélection génomique a démarré en 2009. C'est donc en 2013 que les évaluations classiques risquent d'être impactées pour la première fois.

Il est aussi important d'éviter un biais dans les évaluations classiques nationales car elles sont ensuite utilisées dans les évaluations génétiques internationales. Les index classiques sont dé-regressés pour servir de phénotypes dans les évaluations MACE. Il est donc à craindre qu'un biais dans les évaluations nationales puisse être transmis aux autres populations étrangères et traduit dans les échelles des autres pays. Après une étape de sélection génomique, et si aucune action n'a été mise en place pour la prendre en compte, les évaluations nationales peuvent être biaisées mais elles sont surtout incomplètes au niveau international. L'information manquante sur les candidats éliminés lors de l'étape de sélection génomique, car non transmise, peut générer une autre source d'erreur. Le risque de biais à l'échelle internationale a donc été étudié par simulation sur données réelles. Les évaluations nationales de trois grandes populations Holstein, françaises, allemandes et américaines, ont été utilisées pour simuler les effets de la sélection génomique sur les évaluations internationales : des jeux de données incomplets et éventuellement biaisés ont été générés pour un caractère de production et une intensité de sélection donnée dans les trois populations. Les simulations ont montré que les taureaux issus d'une étape de sélection génomique étaient fortement pénalisés sur l'échelle de

leur pays d'origine comme sur les autres échelles via la matrice de variance-covariance génétique. Le biais généré dans un pays affectait aussi les taureaux contemporains des autres pays via la matrice de parenté. Les classements internationaux étaient nettement perturbés.

En conclusion, une pratique de sélection nouvelle et de grande ampleur comme la sélection génomique doit être prise en compte dans les modèles d'évaluations génétiques pour éviter une série de conséquences néfastes. A l'échelle nationale, les estimations génétiques seraient biaisées et moins précises. Tout biais serait transmis de manière incontrôlée à l'échelle internationale. Ce sont non seulement les classements internationaux et donc le commerce international de semences qui pourraient être impactés mais aussi les futures prédictions génomiques. Une étape de sélection génomique non prise en compte pourrait finalement avoir des conséquences sur l'efficacité globale de la sélection. La mise en place d'une évaluation combinée et itérative permettrait non seulement d'intégrer à court et moyen terme cette nouvelle étape mais aussi de diffuser proprement l'information génomique à toute la population. Les schémas de sélection en bovins laitiers pourraient profiter d'outils de sélection précoce et précis, sans craindre un biais dû à la sélection génomique.

Il reste cependant à développer un test en amont des évaluations internationales pour garantir la qualité des index nationaux vis-à-vis des pratiques de sélection génomique. De plus, les évaluations internationales nécessitent d'être adaptées pour, elles aussi, prendre en compte correctement l'information génomique. Les enjeux sont importants à l'échelle internationale, les évaluations fournies sont aujourd'hui essentielles pour la précision des évaluations génomiques à partir de populations de références multinationales. Gérer de grandes bases de données, de surcroit stratégiques, mettre au point des méthodes d'évaluations satisfaisantes pour tous les acteurs de la sélection dans le monde, à savoir ceux qui ont des outils génomiques et ceux qui n'en ont pas (encore) sont des savoir-faire à développer rapidement. Ce sont de grands défis pour la communauté internationale alors que la collaboration entre centres de calcul peut être remise en question par la compétition entre pays et/ou entreprises de sélection, intensifiée par cette innovation technologique qu'est la sélection génomique. Dans ce nouveau contexte, ce ne sont pas seulement les techniques qui sont en train de changer radicalement mais aussi les rapports de force entre acteurs et de ce fait, l'organisation mondiale de la sélection chez les bovins laitiers.

Mots-clés

Sélection Génomique - Evaluation Génétique – BLUP – MACE – Biais - Bovins Laitiers - Interbull

Acknowledgments/Remerciements

Il y a trois ans, l'INRA et l'UNCEIA m'ont donné l'opportunité de mieux comprendre les « rouages » de la sélection génétique, avec leurs outils et leurs mécaniques. J'ai eu la chance de débuter ma formation doctorale dans « l'ère de la génomique » et de découvrir un monde en pleine révolution, technique comme stratégique. Cette première expérience professionnelle a été d'autant plus riche qu'elle était directement connectée avec la réalité du terrain, les entreprises de sélection, les coopératives d'insémination et les élevages. Interbull m'a fait découvrir une autre dimension, internationale, avec un réseau complexe de collaborations dans un environnement pourtant hautement concurrentiel. A différentes échelles, j'ai ainsi pu assister aux engouements comme aux appréhensions qu'une innovation telle que la génomique peut susciter. Durant ces trois années, j'ai bénéficié d'un contexte de travail particulièrement stimulant et riche en enseignements, techniques comme humains, pour m'aider à compléter le « tableau » que je me faisais du milieu de l'élevage. Ce « tableau », je l'ai débuté très tôt. Initiée par mon père et mon grand-père, j'ai été témoin de leur relation à la terre et aux animaux si particulière. J'ai été témoin de leur passion et leur fierté, si caractéristique des éleveurs, face, pourtant, aux fortes contraintes techniques et économiques du métier. Ce monde complexe mais en évolution constante, a toujours suscité ma curiosité, mon admiration et l'envie de m'y investir, qui n'a été que renforcée par cette expérience de trois ans.

Merci à l'INRA et plus particulièrement au département de génétique animale, unité GABI, sous la direction de Jean-Pierre Bidanel, de m'avoir accueilli à Jouy-en-Josas et fait bénéficier d'un environnement de travail très épanouissant.

Merci à l'UNCEIA, sous la direction de Maurice Barbezant puis Xavier David, pour avoir apporté son support financier et m'avoir aidé à communiquer sur mes travaux de recherche auprès des différents acteurs de la filière génétique en France. Ce lien avec le « terrain » a donné un tout autre sens à mon travail de thèse.

I would like to acknowledge the INTERBULL Centre and its director, João Dürr, for having welcomed me in their team for 5 months, devoting to me time and making available to me tools and data . Also, I would like to especially thank Hossein Jorjani for his supervision. I really appreciated that time in Sweden, discovering another work organization and so many different cultures within only one team!

Je voudrais maintenant remercier, mon directeur de thèse, Vincent Ducrocq, les membres du projet AMASGEN et plus particulièrement ceux de mon comité de thèse, Pascal Croiseau, Sebastien Fritz, Andres Legarra, Sophie Mattalia, Christele Robert et Etienne Verrier, le groupe de travail PROTEJE et le comité de pilotage de la SAM qui ont non seulement suivi l'évolution de mon travail au cours de ses trois années mais surtout aidé à son bon déroulement, de part leurs questions et suggestions.

I would also like to acknowledge all the members of my thesis committee for their review on my manuscript and their interesting questions and comments and especially Zengting Liu and Erling Standberg.

Entre la beauté du BLUP, le côté obscur des QTL et le danger imminent que représente un biais, c'est tout un roman qui aurait pu être écrit ! Cette thèse n'en a pas moins été une riche aventure : les protagonistes sont nombreux et je voudrais maintenant les en remercier personnellement pour avoir enrichi ma propre histoire, au-delà de la transmission des connaissances et du partage d'expériences.

Comment te dire merci, Vincent ? Il y a quelques années, je frappais à la porte de ton bureau pour un stage…en Inde, destination que je n'avais même pas envisagée! Ton immense soutien durant ce stage, malgré la distance et les difficultés propres aux pays, a fortement contribué à mon engagement pour cette thèse, aussi étonnante ait été ma décision au premier abord, moi qui ne te parlais que de « terrain »… Jamais je n'aurais imaginé que je réaliserais ce type de travail : des matrices, des statistiques, d'obscures lignes de code… et des vaches tout de même (mais pas blanches !). C'est bien grâce à toi, à ce temps précieux que tu m'as consacré, et à ton optimisme constant que j'ai pu relever ce défi ! Tu m'as permis d'avoir confiance et surtout de découvrir un monde que j'ignorais totalement – au-delà d'un monde biaisé, bien évidemment ! Je te remercie sincèrement pour cette belle expérience en espérant que tu me soumettras encore à d'autres défis!

Je tiens à remercier Etienne Verrier, de l'AgroParisTech, pour avoir suivi mon parcours depuis déjà de nombreuses années, d'avoir écouté mes questions et de m'avoir encouragée.

Je ne saurais suffisamment remercier les collègues de Jouy et de Bercy pour toutes les riches discussions et échanges qui ont non seulement alimenté mes réflexions sur la thèse (le biais existe-t-il vraiment ?) mais aussi ma connaissance du milieu professionnel et de sa complexité !

Je pense aux discussions tardives mais animées sur le « boulevard » de la génomique - merci à Sébastien et François – mais pas seulement, merci à Sophie Mattalia et Xavier David pour leur partage d'expériences, notamment à l'international, avec Interbull et Eurogenomics.

Je pense aux nombreuses fois où je suis venue toquer à une porte de bureau pour quémander une information, une astuce, ou juste venue pour exposer mes interrogations alors que vous n'aviez rien demandé… Merci pour ce temps partagé et votre soutien, tant professionnel que personnel. Merci à Sophie Allais, ma première victime car souvent voisine, à Rachel pour sa constante bonne humeur qui fait d'elle, une super co-bureau. Merci aussi à Hélène, Armelle, Pascal, Julie, Stéphanie, Sophie Moureaux, Anne, Sandrine pour leur pédagogie et leur grande gentillesse. Merci à l'équipe d'informaticiens pour les 1000 fois où je réclame un mot de passe ! Merci pour les coups de pelle dans le tunnel France-Suède ! Merci aux secrétaires pour tous leurs services !

Classiquement, on enseigne aux thésards, la gestion de projet. A Jouy, je remarque que le dispositif de thèse va encore plus loin, on leur apprend la gestion de l'effort avec des entraineurs sportifs hautement reconnus, enseignant endurance (merci Thierry et Sophie A. pour les foulées du vendredi) et combativité (merci à Eric et Aurélia pour le bad) !

Même si je ne vous cite pas tous : comment cette thèse se serait-elle déroulée sans la pause café et la bonne humeur du 211?

Et surtout, comment la lutte contre le biais se serait-elle terminée sans le « comité spécial de solidarité entre thésards » ??? Avec p'tit Romain qui m'a longuement écouté, Chris, bien solidaire, Romain (merci coloc' pour ton soutien, ton intérêt, et ta capacité à me re-motiver, j'espère pouvoir t'apporter autant d'aide dans ta thèse!) et, et, et… Bérénice, mon ouragan préféré !!! Qu'aurais-je fait sans toi un certain 14 octobre ?! Merci pour toute l'énergie que tu es capable de transmettre ! Je pense aussi à ceux qui ont « quitté le navire » (des thésards seulement !) entre temps : encore une fois, merci ma Soso, pour TOUT !!!, pour toutes les fois où tu m'as écouté, rassuré, fais relativiser !!! Et merci François pour ta précieuse aide au jour le jour ; pour ton immense et constant soutien, au-delà de la thèse et jusqu'en Suède !!!! I will not forget the help of my PhD colleagues in Uppsala, sharing questions but not only: thank you Anne, Mohammad and especially my dear Maria for having made this hard winter so nice!

Mais Bercy n'était pas loin non plus : je voudrais remercier Mr. Barbezant pour ses encouragements, Laurent pour ses commentaires et le regard apporté à ma thèse, merci à Christine pour ses réponses toujours rapides et sa gentillesse, merci à l'ensemble du personnel de l'UNCEIA pour l'accueil fait aux « cifrettes ».

I only spent a few months with you but I felt like in a little family: thank you Hossein, Joao, Jette, Birgitt, Valentina, Eva, Flavio, Dan, Elie for your daily help and very warm welcome, including Fika time!

From this Swedish winter, I would like to thank my "special reviewers" who accepted to read parts of the manuscript and checked that the text was understandable from a linguistic point of view. Many thanks to Olof, Tiago, Jon, Enrique: it must have been hard for you to read about BLUP!

Tout simplement, merci pour votre aide à tous, votre constant soutien, mais aussi pour toutes ces petites conversations qui rendent le quotidien si agréable - un quotidien avec ces fous rires et ces quelques larmes, qui font des collègues, des amis.

Je ne peux pas conclure cette page sans avoir une pensée pour toutes ces personnes qui me sont particulièrement chères et qui, bien qu'impuissantes face au BLUP et au biais, m'ont transmis beaucoup d'énergie au cours de cette thèse et malgré les épreuves de la vie : un immense merci à ma famille INSA, à mes amis de l'Agro, aux amis de Jendin et à ma famille. Tu en fais partie ma Nono, ma Gisèle. Pendant trois ans, au même rythme, nous nous sommes serrées les coudes, mais cela va bien au-delà : merci pour tous ces moments partagés, cette complicité et ta bienveillance. Sans cette énergie dont vous avez toujours fais preuve et que vous m'avez enseigné, Papa, Maman, je n'aurais jamais relevé ce défi. Merci à vous deux.

Je conclurais comme j'ai commencé, par une pensée pour mon père, pensée qui m'aura accompagnée tout au long de la thèse, mon objectif étant de pouvoir continuer à faire vivre ce qu'il m'a transmis…