Innovative Combination of all Sources of Information for Production Traits in Slovenian Brown Swiss

J. Vandenplas[1,2,*], M. Spehar[3,4], K. Potocnik[4], N. Gengler[1] and G. Gorjanc[4,5]
[1] University of Liege, Gembloux Agro-Bio Tech, 5030 Gembloux, Belgium
[2] National Fund for Scientific Research,1000 Brussels, Belgium
[3] Croatian Agricultural Agency, 10000 Zagreb, Croatia
[4] University of Ljubljana, Biotechnical Faculty,1000 Ljubljana, Slovenia
[5] The Roslin Institute and Royal (Dick) School of Veterinary Studies, The University of Edinburgh, Easter Bush, Midlothian, Scotland, UK
[*] Current address: Wageningen UR Livestock Research, Animal Breeding and Genomics Centre, 6700 AH Wageningen. the Netherlands
Email: jeremie.vandenplas@wur.nl

Abstract

Slovenian Brown Swiss is a small population with genetic improvement based on its own breeding program supplemented with imports from other populations. Routine national genetic evaluations for milk, fat, and protein yields are computed from all available national phenotypic and pedigree data. At an international level, a Multiple Across Country Evaluation (MACE) is performed by Interbull to aggregate estimated breeding values (EBV) for international sires across different populations into a single Slovenian ranking. Additionally, a genomic evaluation for many sires is now routinely computed at an international level through the InterGenomics (IG) project. Phenotypes used for this genomic evaluation are deregressed MACE EBV which generate genomically enhanced EBV (GEBV) for all genotyped sires. However, national evaluations are not influenced by these international evaluations and, therefore, may be less accurate and even biased because foreign data used to select foreign sires are not used at the national level. Therefore, an integration of international evaluations back into the national evaluations is required to use the available information in an optimal way for both bulls and cows. The aim of this study was to assess the potential of an innovative Bayesian approach, based on a single-step genomic Best Linear Unbiased Prediction, that combines national data for milk, fat and protein yields with the IG genotypes and information (i.e., GEBV and reliabilities). Because IG information considers genotypes and also MACE information, which also includes national information, double-counting of contributions due to records and due to relationships had to be considered. The integration of IG genotypes and information showed an increase of reliability for the three traits, especially for all IG sires. For example, for IG sires with progeny with national records, the integration led to an average increase of reliability of > 0.10 points for milk yield, in comparison to their average national reliability. For the IG sires without progeny with national records, an average increase of reliability of >0.74 points was observed for the same trait. An average increase of reliability of > 0.05 points was also observed for animals with a reliability <0.30 and sired by genotyped IG sires and with progeny with records. Finally, this approach has the potential to simultaneously combine national data and IG genotypes and information. Furthermore, while it was not implemented in this study, this approach has the advantage to allow the consideration of genotypes of other non-IG animals (e.g., cows).

Key words: single-step genomic BLUP, combination, external, genomically enhanced EBV

Introduction

Simultaneous use of all data by Best Linear Unbiased Prediction is a condition to predict unbiased estimated breeding values (EBV; Henderson, 1984). However, this condition is not always fully met. For example, Slovenian Brown Swiss (BSW) is a small population with genetic improvement based on its own breeding program supplemented with imports from other populations (e.g., less than 10% of the cows are sired by foreign sires), while routine national genetic evaluations for milk, fat, and protein yields are computed only from all available national phenotypic and pedigree data. Indeed, foreign raw data used to select foreign animals is unavailable leading to potential biases in routine national genetic evaluations. National EBV (EBV_N) will be also less reliable because incomplete data is considered. This issue was partially solved for international sires at an international level through the Multiple Across Country Evaluation (MACE; Schaeffer, 1994). Performed by Interbull (Uppsala, Sweden), the MACE aggregates EBV for international sires across populations into a single country‘s specific ranking, e.g., into a Slovenian ranking.

The use of dense marker maps with thousands of single nucleotide polymorphisms (SNPs) and concurrent substantial decrease in genotyping costs (Hayes *et al.,* 2009) have given opportunities to use this information in selection decisions commonly called “genomic selection” (e.g., Meuwissen *et al.,* 2001). SNP genotypes are therefore an additional data source that can improve accuracy of selection and that allows selection of animals early in life (VanRaden, 2008). Unfortunately, the genomic selection could increase these problems for national evaluations. Currently, genomic evaluation systems are mainly based on multi-step approaches using MACE results as a primary source of phenotypic information. In BSW, such a genomic evaluation is performed at an international level through the InterGenomics (IG) project operated at Interbull (Jorjani *et al.,* 2012). The reference population used for this IG genomic evaluation ($EVAL_{IG}$) consisted of almost 8,000 genotyped bulls from seven countries (Austria, France, Germany, Italy, Slovenia, Switzerland, and the United States) and phenotypes are deregressed MACE EBV. Within the scope of the IG consortium, genomically enhanced EBV (GEBV) and associated reliabilities (GREL) for all genotyped BSW sires are generated on a country’s specific scale (Jorjani *et al.*, 2012). A part of the data used for the computation of MACE EBV and, therefore, for GEBV comes from Slovenian population.

Another approach using genomic information is called “single-step genomic evaluation” (ssGBLUP) (Aguilar *et al.,* 2010; Christensen and Lund, 2010). The ssGBLUP could reduce potential biases in the estimation of GEBV by the simultaneous combination of genomic, pedigree and national phenotypic information (VanRaden, 2012), and also because fewer approximations are made than in multi-step approaches. However, in opposition to multi-step genomic approaches, ssGBLUP uses only national information and is unable to directly use other sources of information as, MACE or IG results. Therefore, an integration of international evaluations back into the national evaluations is required to use the available information in an optimal way for both bulls and cows. Because international evaluations, i.e. both MACE and $EVAL_{IG}$, are partially based on national information, their integration could lead to double counting of contributions due to relationships among the international sires or due to records (e.g., if national data is combined with MACE information, partially based on this national data) (Vandenplas and Gengler, 2012; Vandenplas *et al.*, 2014).

The aim of this study was to assess the potential of a Bayesian approach, based on a ssGBLUP, that combines national data for milk, fat, and protein yields with IG genotypes and information (i.e., $GEBV_{IG}$ and reliabilities ($GREL_{IG}$) obtained from $EVAL_{IG}$), by avoiding all possible double-counting. IG information was chosen to be integrated because it is computed from MACE results that already integrate national and foreign information. Such evaluation produced GEBV and GREL on the same scale for sires, cows and young animals. Selection based on such results would be more efficient and transparent.

Materials and Methods

National data, hereafter called “national dataset” included phenotypic data for BSW dairy cattle and was provided by Biotechnical Faculty, Department of Animal Science in April 2014 which is responsible for routine national genetic evaluation (Potočnik, 1999). After edits, it consisted of 1 286 698 test-day records between 5 and 305 days in milk (DIM) for 56 764 BSW cows recorded between years 2000 and 2014 for milk, fat and protein yields. The associated pedigree included 101 522 animals. For the three traits, foreign information consisted in $GEBV_{IG}$ and associated $GREL_{IG}$ routinely computed at an international level by the second routine evaluation in 2014 at Interbull through the IG project. A total of 5 852 genotyped BSW sires (hereafter called "IG sires") were associated with GEBV_{IG} and GREL_{IG} (hereafter called "IG information") for the three studied traits. Also, for each trait, contributions of national information into MACE, and therefore into the IG evaluation, were determined based on the domestic effective daughter equivalents (EDC) reported by Interbull and associated with EBV_{N} and associated REL (REL_{N}). For the three traits, a total of 277 IG sires were associated with a domestic EDC different from 0, meaning that associated EBV_{N} (EBV_{M}) and associated REL_{N} (REL_{M}) contributed to the MACE second routine evaluation in 2014. These EBV_{M} and associated REL_{M} contributing to the MACE routine-run for the 277 IG sires were considered to avoid double counting of contributions due to records, as explained below. Finally, genotypes for the 5 852 IG sires were provided by Interbull. A total of 38 863 SNPs were selected after editing. All information was harmonized between the national evaluation and EVAL_{IG} by adjusting scale and mean difference towards the original expression of the trait in the national genetic evaluation computations.

For the studied traits, the genetic evaluation model used in the national evaluation system is a univariate repeatability test-day model (Potočnik et. al., 2000) and can be written as follows:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_{\text{a}}\mathbf{a} + \mathbf{Z}_{\text{c}}\mathbf{c} + \mathbf{Z}_{\text{p}}\mathbf{p} + \mathbf{e}$$

where **y** is a vector of phenotypic observations for daily milk, fat, or protein yield, **b** is the vector of fixed effects, **a** is the vector of random additive genetic effects, **c** is the vector of random herd effects, **p** is the vector of random permanent environment effects, and **e** is the vector of residuals. The matrices $\mathbf{X}, \mathbf{Z}_{\text{a}}, \mathbf{Z}_{\text{c}}$, and $\mathbf{Z}_{\text{p}}$ are incidence matrices linking **y** and **b**, **a**, **c**, and **p**.

For the national routine evaluation, it was assumed that $\mathbf{a} \sim \text{N}(\mathbf{0}, \mathbf{A}\sigma_a^2)$ where $\mathbf{A}$ is the numerator relationship matrix and σ_a^2 is the additive genetic variance of the considered trait, that $\mathbf{c} \sim \text{N}(\mathbf{0}, \mathbf{I}\sigma_c^2)$ where $\mathbf{I}$ is the identity matrix and σ_c^2 is the variance of herd effect, that $\mathbf{p} \sim \text{N}(\mathbf{0}, \mathbf{I}\sigma_p^2)$ where σ_p^2 is the permanent environment variance, and that $\mathbf{e} \sim \text{N}(\mathbf{0}, \mathbf{I}\sigma_e^2)$ where σ_e^2 is the residual variance.

For the studied traits, two genetic evaluations were performed. First, a national genetic evaluation (EVAL_{N}) was performed similarly to the national routine evaluation, using the national dataset and the associated pedigree. Second, a univariate single-step genomic Bayesian evaluation (EVAL_{B}), as outlined by Vandenplas *et al.* (2014), was performed to combine the national dataset with IG genotypes and IG information available for the 5 852 IG sires. The proposed EVAL_{B} can simultaneously integrate several sources of information into a genetic evaluation and is able to avoid double-counting of contributions due to relationships and due to records. Indeed, as outlined previously, it is worth noting that IG information considers IG genotypes and also MACE information, which also includes national information. The non-consideration of this property of IG information would lead to a double-counting of contributions due to national information and must, therefore, be taken into account. Furthermore, as outlined by Vandenplas *et al.* (2014), the proposed EVAL_{B} can be extended to a ssGBLUP in which the inverse of the numerator relationship matrix ($\mathbf{A^{-1}}$) used in the mixed model equations was replaced by the inverse of the combined genomic-pedigree based relationship matrix ($\mathbf{H^{-1}}$). For this study, $\mathbf{H^{-1}}$ could be written as:

$$\begin{aligned}\mathbf{H^{-1}} &= \begin{bmatrix}\mathbf{H_{N,N}} & \mathbf{H_{N,IG}} \\ \mathbf{H_{IG,N}} & \mathbf{H^{-1}_{IG,IG}}\end{bmatrix}^{-1} \\ &= \mathbf{A^{-1}} + \begin{bmatrix}\mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G^{-1}} - \mathbf{A^{-1}_{IG,IG}}\end{bmatrix}\end{aligned}$$

where the subscripts IG and N refer to the IG sires and to the other non-IG animals (referred by the subscript "N"), respectively, $\mathbf{G^{-1}}$ is the inverse of the genomic relationship matrix (VanRaden, 2008) associated with the IG sires and $\mathbf{A^{-1}_{IG,IG}}$ is the inverse of the numerator relationship matrix for the IG sires.

Therefore, following Vandenplas *et al.* (2014), the combination of the national dataset, IG genotypes and IG information for EVAL_{B} was performed using the following modified mixed model equations:

$$\begin{bmatrix} \mathbf{X'R^{-1}X} & \mathbf{X'R^{-1}Z_a} & \mathbf{X'R^{-1}Z_c} & \mathbf{X'R^{-1}Z_p} \\ \mathbf{Z_a'R^{-1}X} & \mathbf{Z_a'R^{-1}Z_a} + \mathbf{H^{-1}}\sigma_a^2 + \mathbf{\Lambda}_{IG} - \mathbf{\Lambda}_N & \mathbf{Z_a'R^{-1}Z_c} & \mathbf{Z_a'R^{-1}Z_p} \\ \mathbf{Z_c'R^{-1}X} & \mathbf{Z_c'R^{-1}Z_a} & \mathbf{Z_c'R^{-1}Z_c} + \mathbf{I}\sigma_c^2 & \mathbf{Z_c'R^{-1}Z_p} \\ \mathbf{Z_p'R^{-1}X} & \mathbf{Z_p'R^{-1}Z_a} & \mathbf{Z_p'R^{-1}Z_c} & \mathbf{Z_p'R^{-1}Z_p} + \mathbf{I}\sigma_{\mathrm{p}}^{-2} \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{a} \\ \mathbf{c} \\ \mathbf{p} \end{bmatrix} = \begin{bmatrix} \mathbf{X'R^{-1}y} \\ \mathbf{Z_a'R^{-1}y} + \left(\mathbf{H^{-1}}\sigma_a^2 + \mathbf{\Lambda_{IG}}\right)\boldsymbol{\mu}_{\mathbf{IG}} - \left(\mathbf{A^{-1}}\sigma_a^2 + \mathbf{\Lambda_M}\right)\boldsymbol{\mu}_{\mathbf{M}} \\ \mathbf{Z_c'R^{-1}y} \\ \mathbf{Z_p'R^{-1}y} \end{bmatrix}$$

where $\boldsymbol{\mu}_{\mathbf{IG}} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{IG,N}} \\ \boldsymbol{\mu}_{\mathbf{IG,IG}} \end{bmatrix}$ with $\boldsymbol{\mu}_{\mathbf{IG,IG}}$ being the vector of GEBV$_{\mathrm{IG}}$ for the IG sires and $\boldsymbol{\mu}_{\mathbf{IG,N}} = \mathbf{H_{N,IG}H_{IG,IG}^{-1}}\boldsymbol{\mu}_{\mathbf{IG,IG}}$ being the vector of GEBV$_{\mathrm{IG}}$ predicted for the non-IG animals (N), $\boldsymbol{\mu}_{\mathbf{M}} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{NM,N}} \\ \boldsymbol{\mu}_{\mathbf{M,IG}} \end{bmatrix}$ with $\boldsymbol{\mu}_{\mathbf{M,IG}}$ being the vector of EBV$_{\mathrm{M}}$ for the 277 IG sires for which EBV$_{\mathrm{M}}$ and REL$_{\mathrm{M}}$ was considered by EVAL$_{\mathrm{IG}}$ through MACE and $\boldsymbol{\mu}_{\mathbf{M,N}} = \mathbf{A_{N,IG}A_{IG,IG}^{-1}}\boldsymbol{\mu}_{\mathbf{M,IG}}$ being the vector of EBV$_{\mathrm{M}}$ predicted for the non-IG animals (N), and the matrices $\mathbf{\Lambda_{IG}}$ and $\mathbf{\Lambda_M}$ are diagonal covariance matrices associated with IG information and national information considered by EVAL$_{\mathrm{IG}}$, respectively.

The diagonal element of $\mathbf{\Lambda_{IG}}$ for the i$^{\mathrm{th}}$ IG sire was equal to $\mathbf{\Lambda_{IG_{ii}}} = \mathbf{RE_{IG_i}}\sigma_e^{-2}$ where $\mathbf{RE_{IG_i}}$ is equal to the value of records equivalents (RE) only due to own records for IG information. Records equivalents $\mathbf{RE_{IG}}$ were estimated from GREL$_{\mathrm{IG}}$, thanks a two-step algorithm that estimates contributions due to relationships and due to records (Vandenplas and Gengler, 2012) and modified to account for genomic relationships by replacing the pedigree-based relationship matrix ($\mathbf{A_{IG,IG}}$) by the genomic relationship matrix ($\mathbf{G}$) to estimate contributions due genomic relationships. Because IG information was only available for IG sires, diagonal elements associated with non-IG animals were equal to 0. The diagonal elements of $\mathbf{\Lambda_M}$ were computed similarly to those of $\mathbf{\Lambda_{IG}}$. Thereby, only the diagonal elements of $\mathbf{\Lambda_M}$ for the IG sires associated with national information taken into account by MACE, and therefore also by EVAL$_{\mathrm{IG}}$, were equal to $\mathbf{\Lambda_{M_{ii}}} = \mathbf{RE_{M_i}}\sigma_e^{-2}$ where $\mathbf{RE_{M_i}}$ were estimated from REL$_{\mathrm{M}}$ thanks to the two-step algorithm (Vandenplas and Gengler, 2012). For the two evaluations, REL$_{\mathrm{N}}$ for EVAL$_{\mathrm{N}}$ and GREL$_{\mathrm{B}}$ for EVAL$_{\mathrm{B}}$ for the i$^{\mathrm{th}}$ animal were computed as $(G)REL_{xi} = 1 - PEV_{xi}\sigma_a^{-2}$ where the subscript x (x = N, B) referred to EVAL$_{\mathrm{N}}$ or EVAL$_{\mathrm{B}}$ and PEV_{xi} was the prediction error variance associated with the ith animal obtained from the inverse of the LHS of EVAL$_{\mathrm{N}}$ or EVAL$_{\mathrm{B}}$.

For the following comparisons, EVAL$_{\mathrm{IG}}$ was considered as the reference for IG sires because this genomic evaluation combined all available national and foreign information as well as all available genomic information. Comparisons between EVAL$_{\mathrm{IG}}$ and EVAL$_{\mathrm{N}}$ or EVAL$_{\mathrm{B}}$ were performed for (a) IG sires having daughters with records in the national dataset (hereafter called “internally used bulls”; 319 sires) and for (b) IG sires having no daughters with records in the national dataset (hereafter called “internally unused bulls”; 5 533 sires). For the two groups of animals, comparisons between the three evaluations were performed based on: (1) Spearman’s rank correlation coefficients (r_s) of GEBV$_{\mathrm{IG}}$ with EBV$_{\mathrm{N}}$ and GEBV computed from EVAL$_{\mathrm{B}}$ (GEBV$_{\mathrm{B}}$), (2) mean squared errors (MSE) of EBV$_{\mathrm{N}}$ and GEBV$_{\mathrm{B}}$ (i.e., mean squared errors expressed as a percentage of the average MSE of EBV$_{\mathrm{N}}$), (3) regression coefficients (a) and (4) R^2 of the regression of GEBV$_{\mathrm{IG}}$ on EBV$_{\mathrm{N}}$ or GEBV$_{\mathrm{B}}$, and (5) average REL. Effects of integration of IG genotypes and IG information into EVAL$_{\mathrm{N}}$ were studied for (a) Slovenian cows with phenotypes and sired by a IG sire, for (b) for non-Slovenian cows with phenotype and sired by an IG sire (i.e., imported cows), and for (c) Slovenian animals without phenotype and sired by an IG sire. Three groups of Slovenian animals without phenotype and sired by a IG sire were defined depending on their REL$_{\mathrm{N}}$. The first group included Slovenian animals associated with a REL_N lower than 0.50; the second group included animals with a REL_N between 0.50 and 0.75; and the third group included Slovenian animals with a REL_N equal or higher than 0.75. For these groups of animals, effects of integration of IG genotypes and IG information were studied by considering: (1) Spearman's rank correlation coefficients (r_s) between EBV_N and $GEBV_B$, and (2) average REL.

All computations were performed using the BLUPF90 program (Misztal, 2013) modified to implement the Bayesian procedure as proposed by Vandenplas and Gengler (2012) and Vandenplas *et al.* (2014). The used variances were those used for the official national genetic evaluation (e.g., https://rodica.bf.uni-lj.si/web/gov/1502/eval/variance_2.html).

Results & Discussion

Of the 5 852 IG sires associated with IG information, a total of 277 IG sires were also associated with national information that contributed to MACE, and thereby also to the IG evaluation. On average, each of the 319 internally used sires had 134.9 daughters with Slovenian records for the three traits.

Internally used bulls

For the 319 internally used bulls, the average $GREL_{IG}$ (SD) was equal to 0.97 (0.02), and the average REL_N (SD) was equal to 0.87 (0.19) for the three traits. Results for r, MSE, a, and R^2 for prediction of $GEBV_{IG}$ by $EVAL_N$ and $EVAL_B$ are shown for the internally used bulls in Table 1 for the three traits. Consideration of IG genotypes and IG information by taking into account double-counting of contributions due to pedigree and genomic relationships and due to records (i.e. $EVAL_B$) led to a ranking that was more similar to the IG ranking (e.g., $r_s > 0.99$ for milk yield), in comparison to the ranking of $EVAL_N$ (e.g., $r_s = 0.79$ for milk yield) for the three traits, even if these sires had already a high number of progeny with Slovenian records. The MSE, a, R^2 and REL showed that accuracy of predictions of $GEBV_{IG}$ by EBV_N or $GEBV_B$ increased when IG genotypes and IG information was integrated (Table 1). For example, the average REL increased by 0.10 points for milk yield, corresponding to an increase of 365.6 daughter equivalents (DE).

Table 1. Parameters obtained between the InterGenomics evaluation ($EVAL_{IG}$) and the national ($EVAL_N$) and Bayesian evaluations ($EVAL_B$) for the 319 internally used bulls.

Trait	Genetic evaluation	r_s [1]	MSE[2]	a[3]	R^2 [3]	REL[4]
Milk yield	$EVAL_N$	0.79	100.00	0.92±0.04	0.68	0.87(0.19)
	$EVAL_B$	>0.99	1.13	1.00±0.003	0.99	0.97(0.02)
Fat yield	$EVAL_N$	0.76	100.00	0.90±0.04	0.62	0.87(0.20)
	$EVAL_B$	0.99	0.88	1.01±0.003	0.99	0.97(0.02)
Protein yield	$EVAL_N$	0.76	100.00	0.90±0.04	0.62	0.87(0.19)
	$EVAL_B$	0.99	10.53	1.00±0.003	0.99	0.97(0.02)

[1] r_s = rank correlation between $EVAL_{IG}$ and $EVAL_N$, or $EVAL_B$;
[2] MSE = mean squared error expressed as a percentage of the average national mean squared error;
[3] a = regression coefficient (SE in parentheses) and R^2 = coefficient of determination of the regression of $GEBV_{IG}$ on (G)EBV estimated by $EVAL_N$, or $EVAL_B$;
[4] REL = average reliability (SD in parentheses).

Internally unused bulls

For the 5 533 internally unused bulls, the average $GREL_{IG}$ (SD) was equal to 0.91 (0.02), and the average REL_N (SD) was equal to 0.17 (0.10) for the studied traits. Results for r_s, MSE, a, and R^2 for prediction of $GEBV_{IG}$ by $EVAL_N$ and $EVAL_B$ are shown for the internally unused bulls in Table 2 for the studied traits. Consideration of IG genotypes and IG information by the Bayesian approach (i.e. $EVAL_B$) led to a ranking that was more similar to the IG ranking (e.g., $r_s > 0.99$ for milk yield), in comparison to the ranking of $EVAL_N$ (e.g., $r_s = 0.55$ for milk yield) for the three traits. The MSE, a, R^2 and REL showed that accuracy of predictions of $GEBV_{IG}$ by EBV_N or $GEBV_B$ increased when IG genotypes and IG information was integrated (Table 2). For example, the average REL increased by 0.74 points for milk yield, corresponding to an increase of 141.24 DE. These results are very interesting because $EVAL_B$ would allow the Slovenian breeders to select more accurately bulls to mate with their animals in a greater range of (foreign) bulls more accurately evaluated on the same Slovenian scale.

Non-IG animals

A total of 43 007 Slovenian cows sired by a IG sire were also associated with phenotypes for the three studied traits. As expected, integration of IG genotypes and IG information into $EVAL_N$ had little effects on their evaluation. The rank correlation between EBV_N and $GEBV_B$ was higher than 0.99.

Of the non-Slovenian animals, 23 cows were sired by an internally used bull and were associated with Slovenian phenotypes. Integration of IG genotypes and IG information into $EVAL_N$ for their sires had little effects on their evaluation, as shown by the parameters r_s, a, and R^2 (Table 3). However, integration increased by at least 4% the average REL associated with these non-Slovenian cows (Table 3).

Table 2. Parameters obtained between the InterGenomics evaluation ($EVAL_{IG}$) and the national ($EVAL_N$) and Bayesian evaluations ($EVAL_B$) for the 5 533 internally unused bulls.

Trait	Genetic evaluation	r_s[1]	MSE[2]	a[3]	R^2 [3]	REL[4]
Milk yield	$EVAL_N$	0.55	100.00	1.40±0.03	0.35	0.17(0.10)
	$EVAL_B$	>0.99	0.04	1.00±0.000	>0.99	0.91(0.02)
Fat yield	$EVAL_N$	0.51	100.00	1.41±0.03	0.33	0.17(0.10)
	$EVAL_B$	>0.99	0.03	1.00±0.000	>0.99	0.90(0.02)
Protein yield	$EVAL_N$	0.60	100.00	1.53±0.02	0.41	0.17(0.10)
	$EVAL_B$	>0.99	0.05	1.00±0.000	>0.99	0.90(0.02)

[1] r_s = rank correlation between $EVAL_{IG}$ and $EVAL_N$, or $EVAL_B$;
[2] MSE = mean squared error expressed as a percentage of the average national mean squared error;
[3] a = regression coefficient (SE in parentheses) and R^2 = coefficient of determination of the regression of $GEBV_{IG}$ on (G)EBV estimated by $EVAL_{N,}$ or $EVAL_B$;
[4] REL = average reliability (SD in parentheses).

Table 3. Parameters for 23 non-Slovenian animals with Slovenian phenotypes and sired by an internally used bull between the national evaluation ($EVAL_N$) and the Bayesian evaluation ($EVAL_B$).

Trait	r_s[1]	REL_N[2]	$GREL_B$[3]
Milk yield	0.98	0.67 (0.14)	0.70(0.12)
Fat yield	0.98	0.64 (0.16)	0.67(0.13)
Protein yield	0.99	0.67 (0.15)	0.70(0.12)

[1] r_s = rank correlation between $EVAL_N$ and $EVAL_B$;
[2] REL_N = average reliability (SD in parentheses) for $EVAL_N$;
[3] $GREL_B$ = average reliability (SD in parentheses) for $EVAL_B$.

Of the Slovenian animals without phenotype and sired by a IG sire, about 335 animals were associated with a REL_N between 0.50 and 0.75 and about 100 animals were associated with a REL_N higher than 0.75, for the three traits (Table 4). Because EBV_N of these animals were already moderately to highly accurate, integration of IG genotypes and IG information into $EVAL_N$ for their sires had little effect on their evaluation, following the different parameters r_s, a, R^2, REL_N and $GREL_B$ reported in the Table 4. It is worth noting that these two groups included, e.g., Slovenian sires having daughters with milk records. The last of the three groups of Slovenian animals without phenotype and sired by a IG sire had a REL_N lower than 0.50 and included about 1 535 animals. This group of animals is interesting to study because it included, e.g., young animals evaluated with a low REL on which a selection decision should be performed. Higher accurate evaluations for this group of animals could therefore increase the accuracy and timeliness of selection. Integration of IG genotypes and IG information into $EVAL_N$ for their sires led to a re-ranking of these Slovenian animals without phenotype in comparison to the ranking of $EVAL_N$ for the three traits (e.g., r_s = 0.95 for milk yield). An average increase of REL of >0.01 points was observed for this group of Slovenian animals without phenotype, for the studied traits. However, this average increase was higher than 0.05 points for Slovenian animals without phenotype and with a REL_N lower than 0.30 (Table 4; Figure 1).

Table 4. Parameters obtained between the national evaluation ($EVAL_N$) and the Bayesian evaluation ($EVAL_B$) for Slovenian animals without phenotype and sired by an internally used bull.

Trait	Range[1]	N[2]	r_s[3]	REL_N[4]	$GREL_B$[5]
Milk yield	]0.00 – 0.50[	1 520	0.95	0.40 (0.07)	0.41(0.06)
	[0.50 – 0.75[	348	0.99	0.57 (0.07)	0.58(0.07)
	[0.75 – 0.99]	103	0.99	0.85 (0.08)	0.85(0.08)
Fat yield	]0.00 – 0.50[	1 571	0.96	0.40 (0.07)	0.41(0.06)
	[0.50 – 0.75[	307	0.99	0.58 (0.07)	0.59(0.08)
	[0.75 – 0.99]	93	0.99	0.85 (0.08)	0.85(0.08)
Protein yield	]0.00 – 0.50[	1 516	0.96	0.40 (0.07)	0.41(0.06)
	[0.50 – 0.75[	352	0.99	0.57 (0.07)	0.58(0.07)
	[0.75 – 0.99]	103	0.99	0.85 (0.08)	0.85(0.08)

[1] Range = Range of REL_N[5];
[2] N = Number of Slovenian animals without phenotype and sired by internally used sires;
[3] r_s = rank correlation between $EVAL_N$ and $EVAL_B$;
[4] REL_N = average reliability (SD in parentheses) for $EVAL_N$;
[5] $GREL_B$ = average reliability (SD in parentheses) for $EVAL_B$.

To our knowledge, integration of genomic-related information into a ssGBLUP was not performed yet. However, the potential of the Bayesian approach was previously studied for blending MACE information with national information into a ssGBLUP to develop a Walloon genomic evaluation in Belgium (Colinet *et al.*, 2013). The developed Walloon genomic evaluation system combined successfully MACE and national information into a ssGBLUP and was successfully tested inside the GEBV tests of Interbull for several traits. For our study, results showed that the Bayesian approach has also the potential to simultaneously combine national data, IG genotypes, and $GEBV_{IG}$ and associated $GREL_{IG}$ into a ssGBLUP. Furthermore, the Bayesian approach can take into account double counting of contributions due to genomic and pedigree-based relationships and due to own records thanks to a two-step algorithm (Vandenplas and Gengler, 2012). Integration of IG genotypes and IG information led to more accurate predictions for the genotyped animals, especially for those without Slovenian daughters with phenotypes. IG information was also propagated to the progeny of the IG sires, which could be useful to take more accurate selection decisions for young animals. Finally, another advantage of this proposed Bayesian approach is that it allows the consideration of genotypes of other non-IG evaluated animals (e.g., cows, young animals), to have estimated BV in the same scale as sires.

Conclusions

In conclusion, the proposed Bayesian approach has the potential to simultaneously combine national data and IG genotypes and IG information into a ssGBLUP. Furthermore, while it was not implemented in this study, the Bayesian approach has the advantage to allow the consideration of genotypes of other non-IG animals (e.g., young animals, cows) by including them in the evaluation.

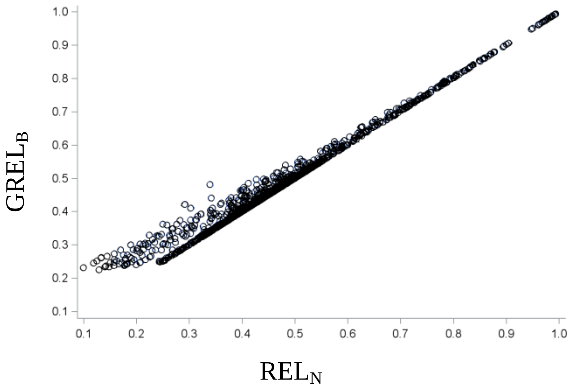

Figure 1. Reliabilities computed from the Bayesian evaluation ($GREL_B$) versus reliabilities computed from the national evaluation (REL_N) for Slovenian animals without phenotype and sired by an internally used sire for milk yield.

Acknowledgements

Jérémie Vandenplas (as a research fellow) and Nicolas Gengler (as a former senior research associate) acknowledge the support of the National Fund for Scientific Research (Brussels, Belgium) for their positions. Gregor Gorjanc, Klemen Potočnik and Marija Špehar acknowledge the Slovenian Ministry of Agriculture, Forestry and Food for financial support through ARRS-CRP-ZP-2012-05/48 project in genotyping of BSW bulls. We also acknowledge the InterGenomics Management Committee for providing genotypes for BSW bulls.

References

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

Christensen, O.F. & Lund, M.S. 2010. Genomic prediction when some animals are not genotyped. *Genet. Sel. Evol. 42*:2.

Colinet, F.G., Vandenplas, J., Faux, P., Vanderick, S., Renaville, R., Bertozzi, C., Hubin, X. & Gengler, N. 2013. Walloon single-step genomic evaluation system integrating local and MACE EBV. *Interbull Bulletin 47*, 203–210.

Jorjani H., Jakobsen J., Hjerpe E., Palucci V. & Dürr, J. 2012. Status of genomic evaluation in the Brown Swiss populations. *Interbull Bulletin 46*, 46–54.

Meuwissen, T.H.E., Hayes, B.J. & Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics 157*, 1819–1829.

Misztal, I. 2013. BLUPF90 family of programs. University of Georgia, Athens. Accessed Jan. 15, 2013. http://nce.ads.uga.edu/wiki/doku.php

Potočnik, K. 1999. Napovedi plemenskih vrednosti za lastnosti mlecnosti pri govedu na podlagi dnevne kontrole. (Prediction of breeding values for dairy production traits using test day model). Master's thesis. Ljubljana, University of Ljubljana, Biotechnical Faculty, Zootehnical Department: 90 p.

Potočnik, K., Kovač, M. & Pogačar, J. 2000. Estimation of dispersion parameters within an animal test day model for Slovenian black and white breed. Zbornik Biotehniške fakultete Univerze v Ljubljani, Kmetijstvo, Zootehnika, Agricultural issue, Zootechnica 76(1), 49-60.

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

Vandenplas, J., Colinet, F.G. & Gengler, N. 2014. Unified method to integrate and blend several, potentially related, sources of information for genetic evaluation. *Genet. Sel. Evol. 46*:59.

Vandenplas, J. & Gengler, N. 2012. Comparison and improvements of different Bayesian procedures to integrate external information into genetic evaluations. *J. Dairy Sci. 95*, 1513–1526.

VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci. 91*, 4414–4423.