INTERBULL BULLETIN NO. 61. 21-22 June 2025, Louisville, Kentucky, USA

Implementation of a Single-Step genomic evaluation system for dairy cattle in Wallonia, Belgium

S. Vanderick*[1*], K. Wijnrocx*[1*]***, J. Vandenplas*[2]***, A. Gillon*[3]*** and N. Gengler*[1]

[1] Animal Sciences, Terra Teaching and Research Centre, University of Liège-Gembloux Agro-Bio Tech, 5030 Gembloux, Belgium

[2] Animal Breeding and Genomics, Wageningen University and Research, 6708PB Wageningen, the Netherlands

[3] Elevéo asbl, AWE Groupe, 5590 Ciney, Belgium

Corresponding author: Katrien.Wijnrocx@uliege.be

** Both authors contributed equally*

Abstract

For several years, dairy cattle breeders in the Walloon Region of Belgium have had access to locally estimated breeding values (EBV) for traits of interest. These evaluations enable Wallonia to contribute to the Multiple Across Countries Evaluation (MACE) conducted by Interbull. In the current local genomic evaluation framework, genomic and pedigree data are integrated with local EBV and external information, MACE-derived EBV (MACE EBV), through a pseudo-single-step genomic evaluation system, producing genomically enhanced EBV (GEBV). However, this approach may introduce biases. To address this, the present study aimed to implement a single-step genomic BLUP (ssGBLUP) that simultaneously incorporates all available national data alongside MACE information and validate this method using milk production traits. The proposed strategy first involves defining "pseudo-traits" that represent MACE traits (i.e., 305-day averages for milk, fat, and protein yields over the first three lactations). MACE EBV are then transformed into adjusted pseudo-phenotypes (i.e., deregressed proofs) and effective contributions but avoiding double-counting of Walloon data within MACE EBV. Next, the variance-covariance matrices from the local random regression test-day model were modified to include the three MACE pseudo-traits as correlated traits. Finally, a single-step genomic evaluation was performed, jointly analyzing test-day records and MACE pseudo-phenotypes. Validation of both pedigree-based and single-step genomic evaluations, both integrating MACE information, was carried out using data from the official Walloon genetic evaluations of April 2022. Results show that MACE information is adequately integrated in the local evaluations, because Pearson correlations between MACE EBV and the integrated EBV were higher than 0.97 across traits. The addition of genomic information in single-step evaluations resulted in small changes for all individuals, as illustrated by Pearson correlations ranging from 0.975 and 0.986 for sires with MACE information.

Key words: dairy cattle, single-step GBLUP, multiple across-country evaluation, milk production traits

Introduction

Walloon Holstein dairy cattle breeders are provided with locally estimated breeding values (EBV) for production-, conformation-, udder health and other functional traits. Nearly all traits are submitted to Interbull (https://interbull.org, Uppsala, Sweden), which performs an across-country genetic evaluation using a multiple across-country evaluation (MACE) system, resulting in MACE-derived international EBV hereafter called MACE EBV (Schaeffer, 1994).

In the current Walloon genomic evaluation system, these MACE EBV are combined with pedigree and genomic information in a pseudo single-step evaluation using a Bayesian integration procedure, accounting for double-counting, to produce local genomically enhanced EBV (GEBV) (Vandenplas et al., 2014; Colinet et al., 2018). This system enables the integration of external (international) information across all animals and datasets in the local system, ultimately enhancing selection decisions (Vandenplas et al., 2015; Vandenplas et al., 2017; Bonifazi et al., 2023). Moreover, post-processing steps are performed to integrate MACE, GMACE and local EBV/GEBV before publication.

However, this pseudo single-step evaluation may introduce biases. In particular, this approach still relies on a first BLUP step which does not estimate fixed effects accounting correctly for genetic differences as reflected by final GEBV. To address this, single-step GBLUP (ssGBLUP) is the method of choice as it combines phenotypic data, pedigree and genomic information simultaneously, resulting in GEBV that are both more accurate and less biased (Misztal et al. 2009, Aguilar et al, 2010, Christensen and Lund, 2010).

The objectives of this study were to develop and validate a single-step genomic evaluation that simultaneously incorporates all available local data alongside MACE information for milk production traits in Walloon Holstein dairy cattle. Integration of MACE information in the current system is very important and had to be conserved. This study is part of an ongoing effort to implement this strategy for all our current evaluations but also for novel traits such as enteric methane emissions.

Materials and Methods

Data

Phenotypic data

All data were provided by the Walloon breeding association Elevéo (Awé Groupe, Ciney, Belgium), which manages performance recording data in the Walloon region of Belgium. Data used in the study were from the official Walloon routine April 2022 genetic evaluations.

Three types of data were used: 1) 4 851 501 test-day records for 305-day milk (MY), fat (FY), and protein yields (PY) across three lactations, 2) MACE EBV and reliabilities of 12 547 bulls obtained from Interbull for these traits, and 3) local EBV and reliabilities of 2 230 local Walloon bulls that were sent to Interbull.

Furthermore, the cleaned pedigree used for the April 2022 routine evaluation consisted of 4851501 animals. Genetic groups were defined as in the current evaluations by group of birth years, origin (Europe vs. USA) and sex.

Genomic data

Genomic data were available for 13 604 animals and consisted of 30 554 single nucleotide polymorphisms (SNPs), routinely used in the Walloon genetic evaluation system. Genotyping was done using the BovineSNP50 Beadchip v1 to v3 and EuroG MD (SI) v9 (Illumina) chips. SNPs common across the four chips were retained, while non-mapped SNPs, SNPs located on sex chromosomes, and triallelic SNPs were excluded. SNPs exhibiting Mendelian conflicts or with a minor allele frequency less than 5% were excluded. The difference between observed and expected heterozygosity was estimated, and SNPs with a difference greater than 0.15 were excluded (Wiggans et al., 2009). After applying all quality control measures, non-mapped SNP, SNP located on sex chromosomes, SNP with Mendelian conflicts, and those with minor allele frequency less than 5% were excluded. Finally, data of 28 470 SNPs located on 29 chromosomes were used.

Definition of MACE Pseudo-traits

Adding external information from MACE discounting for already contained local information has been a topic of applied research for a long time (Gengler & Vanderick, 2008). Recently, a commonly accepted strategy was developed that we also implemented. Following the framework of Vandenplas et al. (2015), MACE EBV were transformed into deregressed proofs (DRP), i.e., “pseudo-phenotypes”, for MY, FY, PY for MACE bulls excluding local information. These pseudo-traits captured the genetic signal embedded in MACE EBV, while avoiding double-counting of national records. They were associated to weights called effective record contributions (ERC), expressing how many record equivalences would have generated the same information content.

Following the strategy developed by Bonifazi et al. (2023), deregression was performed using the method of Garrick et al. (2009) where EBV were adjusted by their reliability to produce DRP. There are two inherent issues that Bonifazi et al. (2023) solved: the elimination of local information in MACE EBV, and the deregression applied to EBV corrected for parent averages (PA), i.e. Mendelian Sampling deviation. Reliability (REL) of this deviation is needed in the deregession process. This value can be obtained by transforming REL to ERC and subtracting ERC associated with PA REL from total ERC. This yields dERC, necessary to properly weigh the DRP used as pseudo-phenotypes in an evaluation. In a final step double counting is eliminated by subtracting local information sent to Interbull generating DRP* and dERC* free of local information.

Models and analysis

The model currently used in the Walloon region of Belgium is a three-lactation, three-trait random regression test-day model (RRTDM) for milk, fat and protein yields. Average lactation EBV are generated computing weighted sums of the underlying genetic random regression solutions resulting in aggregate 305-DIM three-lactation EBV that are sent to Interbull (EBV_sent). Details can be found in Auvray & Gengler (2002) and Croquet et al. (2006).

We developed the new model to be as close as possible to the current Walloon genomic evaluation, which uses a pseudo-single-step evaluation as outlined in Vandenplas et al. (2014) and Colinet et al. (2018). Briefly explained, local EBV and MACE EBV were combined in a system using a combined pedigree and genomic relationships matrix, while discounting for local information included in MACE (EBV_sent), resulting in local GEBV. An overview of this system can be found in Figure 1A.

The new developed system relies on the simultaneous analysis of phenotypic data, instead of local EBV, the MACE pseudo-phenotypes (as DRP) weighted by ERC, and pedigree and genomic information using ssGBLUP (Aguilar et al., 2010), resulting in GEBV (GEBV_ss).

Another problem is that the local traits and MACE traits are not on the same scale. Recently some research (e.g. Boerner et al., 2023) has proposed complex solutions. The complication is, however, only because one tries to pass from a 305DIM 3-lactation MACE trait represented by a single EBV to a large number of genetic effects (i.e., 3 random regressions) and traits (i.e. 3 lactations). Complicated and potentially imprecise backsolving equations can be avoided by using the following process. The variance-covariance matrices from the national test-day random regression model were modified to include the three MACE pseudo-traits as correlated traits but initially without records linking these to the Legendre polynomials across the 3 lactations. As the linear function in the direct local evaluation solutions to MACE EBV is a simple transformation, transforming by the same function initial variance-covariance matrices generates the needed augmented genetic (co)variance matrices (Vandenplas et al., 2015). Its singularity was avoided by multiplying the used (co)variance between random regressions and MACE traits by a factor of 0.999. An interesting side effect of this reparameterization is that the new system of equation generates natively EBV to be sent to the MACE runs of Interbull. Moreover, the setting of this equation system is very flexible, so, one can choose to add or not MACE bulls. The system of equations can also be easily modified to keep fixed effects constant. This will allow us to provide Interbull with the different requested

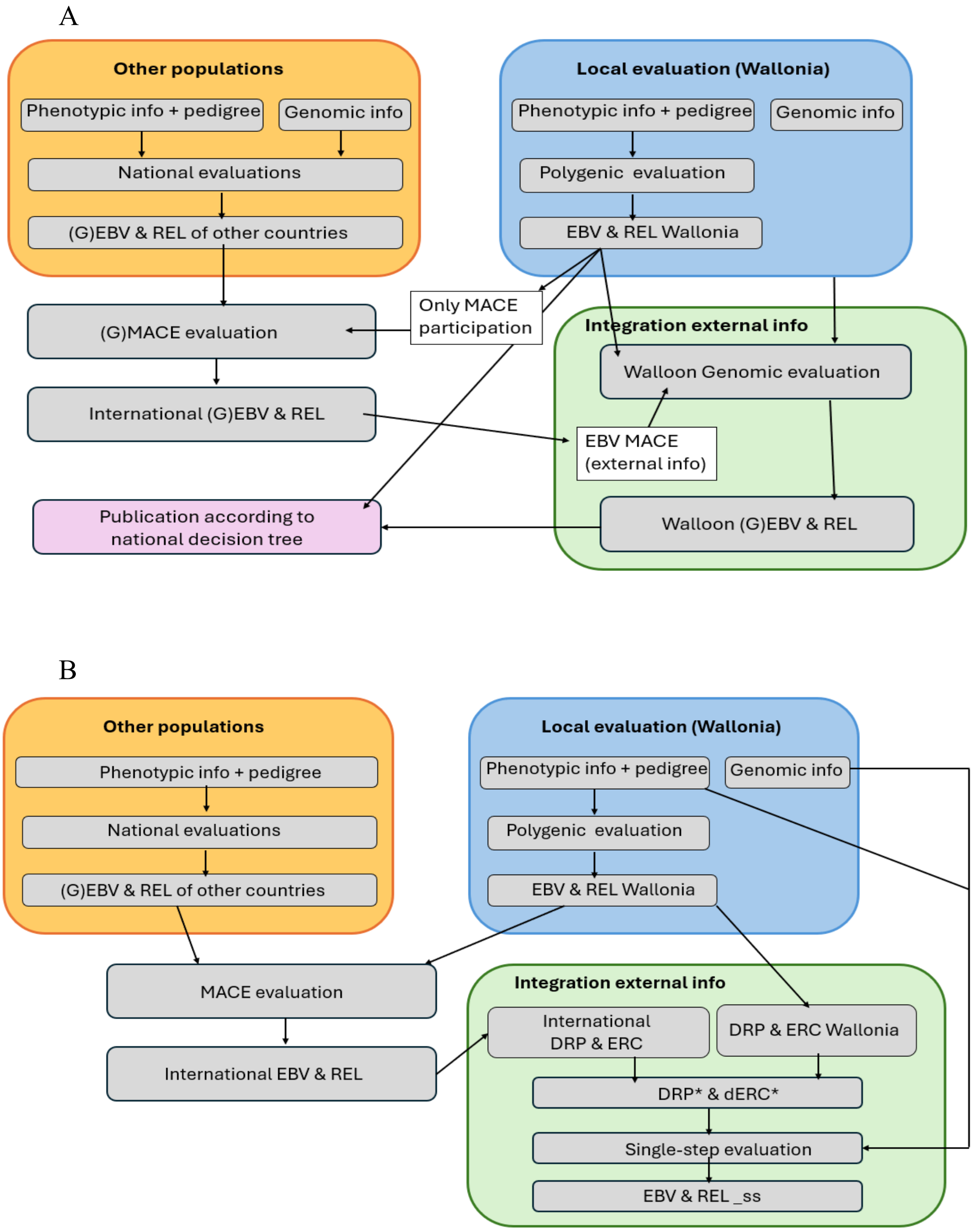

Figure 1. Comparison of the currently applied genomic evaluation system in Wallonia (above, A), compared to the developed single-step system with integration of MACE information based on DRP (below, B).

unbiased EBV in the future (i.e., computing EBV_sent but using fixed effects that are obtained by solving with **H**).

Figure 1B gives an overview of the developed single-step evaluation system with incorporation of MACE data based on DRP* and dERC*. Software used for these calculations was based on BLUPF90 family of programs (Misztal et al., 2014) but with a certain number of adaptations including checking of residuals for outliers and other adaptations to improve its usefulness in the context of genetic evaluations.

Furthermore, the calculation of genomic reliabilities had to be adapted as well. First, pedigree reliability (REL) has been determined using Effective Daughter Contributions and following standard procedures employed in the current genetic evaluations for yield traits in the Walloon Region of Belgium, as described by Strandén et al. (2000), but extended to allow integration of ERC of MACE bulls. Second, genomic reliability (GREL) was calculated using an approach adapted from Gao et al. (2023), using the previously obtained pedigree REL as a starting value after adjustment for double-counting due to pedigree information (Zaabza et al., 2022).

Validation

To assess consistency, we did two types of comparisons. First, to validate the integration of MACE information in a genetic evaluation, we compared for the same bulls MACE EBV with EBV_ss obtained from a pedigree-based evaluation integrating MACE information as in the developed ssGBLUP. Theoretically, the system should yield exactly the same EBV_ss to MACE EBV. Then, the GEBV from the current genomic evaluation system were compared to GEBV_ss obtained using the developed ssGBLUP system (GEBV_ss), both systems integrating EBV MACE. In this study we compared correlations and regression coefficients, the first establishing the relationship between sets of breeding values, the later assessing potential changes in variances (i.e., inflation of deflation) of the sets of breeding values that are compared.

Results

Validation of integration of MACE information in the genetic evaluation

Tables 1 and 2 present Pearson correlations among EBV obtained from the current local pedigree-based evaluation (EBV_sent), MACE EBV, EBV obtained from the pedigree-based evaluation integrating MACE EBV (EBV_ss), and GEBV obtained from the new ssGBLUP with integration of MACE EBV (GEBV_ss). Table 1 specifically shows the correlations for milk, fat and protein yields for all bulls included in the MACE evaluation, as well as for own local bulls participating in MACE.

Table 1: Pearson correlations between (G)EBV_ss with MACE EBV and EBV_sent, for MACE bulls, with or without any local information sent to INTERBULL, for milk (MY), fat (FY) and protein (PY) yields.

		MACE EBV (N=12547)	EBV_SENT (N=2230)
MY	EBV_ss	0.967	0.962
	GEBV_ss	0.969	0.960
FY	EBV_ss	0.945	0.958
	GEBV_ss	0.950	0.955
PY	EBV_ss	0.979	0.964
	GEBV_ss	0.980	0.961

Table 2 shows the same correlations, but only for genotyped bulls.

Table 2: Pearson correlations between (G)EBV_ss with MACE EBV and EBV_SENT, for local genotyped MACE sires, with or without any local information sent to INTERBULL, for milk MY), fat (FY) and protein yields (PY).

		MACE EBV (N=2828)	EBV_SENT (N=1036)
MY	EBV_ss	0.965	0.961
	GEBV_ss	0.973	0.956
FY	EBV_ss	0.951	0.961
	GEBV_ss	0.967	0.955
PY	EBV_ss	0.979	0.967
	GEBV_ss	0.979	0.962

The developed approach enabled integration of MACE EBV into multi-trait random regression test-day evaluations. This is shown by correlations between MACE EBV and EBV_ss that range between 0.95 and 0.98, and by regression coefficients close to 1 (Figure 2).

The new ssGBLUP approach enabled multi-trait random regression test-day ssGBLUP with integration of MACE EBV. The impact of genomic correlations on the evaluation of genotyped bulls is shown by the Pearson correlations between MACE EBV and GEBV_ss, ranging from 0.945 to 0.980. Compared to EBV_ss, we noticed a very small augmentation of correlations when genomic information is integrated (Table 1), and this augmentation is more visible when only looking at genotyped bulls (Table 2). Additionally, Figure 3 shows again excellent R² values and regression coefficients close to 1 (therefore no bias) with inclusion of genomic information, indicating a perfectly valid integration. Figures for fat and protein yields are not displayed but show the same tendency.

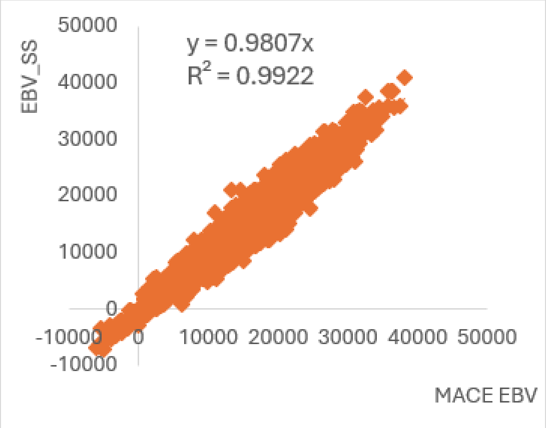

Figure 2. EBV_ss versus MACE EBV for all MACE bulls for milk yield (x10) expressed on the same scale

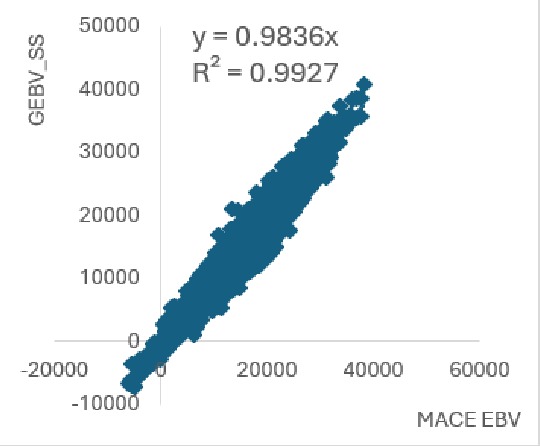

Figure 3. GEBV_ss versus MACE EBV for all MACE bulls for milk yield (x10) expressed on the same scales

Validation of the integration of genomic information

Walloon GEBV obtained from the current and developed genomic systems are similar. This is shown by Pearson correlations between GEBV calculated with the current and developed systems of 0.975 (fat yield), 0.984 (milk yield) and 0.986 (protein yield) for the 2828 genotyped MACE bulls. Additionally, Figure 4 shows again a very high R² value. Graphics for fat and protein yields are not displayed but show the same tendency.

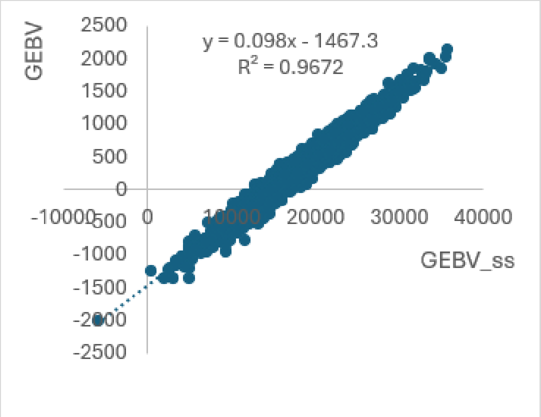

Figure 4. GEBV computed with the current (GEBV) and the developed (GEBV_ss) (x10) genomic system for all genotyped MACE bulls for milk yield, not expressed on the same scales.

Discussion

This study aimed to validate a single-step analysis that simultaneously incorporates all Walloon national data with available MACE information for milk production traits in Walloon Holstein dairy cattle. In recent years, ssGBLUP has become the preferred approach for genomic evaluation. Multiple countries are shifting their evaluations towards single-step evaluations, which has been reported to demonstrate less bias and higher accuracies, as proven by several countries in dairy cattle, (Zavadilova et al., 2014; Mäntysaari et al., 2020; Alkhoder et al, 2022; Himmelbauer et al., 2021; Pimentel et al. 2021, Cesarani et al., 2021); Guarini et al., 2019)) amongst others. Furthermore, studies in beef cattle (e.g., Lourenco et al. 2015, Bonifazi et al. 2023) have also demonstrated effectiveness. Our results show very high correlations and regression coefficients close to one, indicating good robustness and thus validating the method.

In the next step towards practical implementation, REL calculation will be performed according to the proposed strategy hereabove. This strategy has already been validated in our local methane evaluation (Chen et al., 2025) but has yet to be validated in the developed Walloon routine evaluation.

The deregression phase of MACE information is based on the quality of the parent averages, calculated directly by Interbull using a sire-grand-sire model. So, for a foreign bull with little or no information in Wallonia, the information is less reliable. However, as expected, as the number of daughters for this bull in Wallonia increases, most of the information in the bull proof comes from the daughters and/or sons (Mrode & Swanson, 1999). So, in the future, the quality of the DRP is expected to further improve.

Conclusions

As a conclusion, the proposed single-step GBLUP enabled an efficient genomic evaluation process, with robust validation statistics, highlighting its feasibility for routine use in Wallonia’s genomic evaluations. The comparison between the current and developed genomic systems ensures that base change modifications of (G)EBV are acceptable, and that limited changes will be experienced by farmers. Future developments will continue to improve the evaluation system, such as for example a re-estimation of the variance components.

Acknowledgments

The authors would like to acknowledge the Walloon Breeders Association (Elevéo) for the data provision and endless support throughout the past years. Furthermore, we would like to thank the Walloon Region (SPW Wallonie) for its financial support, and ULiège for financing a travel grant.

References

Aguilar, I.; Misztal, I.; Johnson, D.; Legarra, A.; Tsuruta, S.; Lawlor, T. 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.

Auvray, B and Gengler, N. 2002. Feasibility of a Walloon Test-Day Model and Study of its Potential as Tool for Selection and Management. In *Proceedings of the 2002 Interbull meeting*. Interlaaken, Switzerland. 123-127.

Alkhoder, H., Z. Liu, D. Segelke, R. Reents 2022. Comparison of a single-step with a multistep single nucleotide polymorphism best linear unbiased predictor model for genomic evaluation of conformation traits in

German Holsteins. *J Dairy Sci.*, Volume 105, Issue 4, 3306 - 3322

Boerner, V., T.T. Nguyen, and G.J. Nieuwhof. 2023. Integration of Interbull's multiple across-country evaluation approach breeding values into the multiple-trait single-step random regression test-day genetic evaluation for yield traits of Australian Red breeds. *J Dairy Sci.* 106(2):1159-1167.

Bonifazi, R. Mario P. L. Calus, Jan ten Napel, Roel F. Veerkamp, Stefano Biffani, Martino Cassandro, Simone Savoia and Jérémie Vandenplas. 2023. Integration of beef cattle international pedigree and genomic estimated breeding values into national evaluations, with an application to the Italian Limousin population. *Genetics Selection Evolution* , 55:41.

Cesarani, A, Masuda, Y, Tsuruta, S, Nicolazzi, E.L., VanRaden, P.M., Lourenco, D and Misztal, I. 2021. Genomic predictions for yield traits in US Holsteins with unknown parent groups. *J. Dairy Sci.*, 104, pp. 5843-5853.

Chen, Y, Atashi, H., Wijnrocx, K., Vandewalle, M., Berat, H., Vanlierde, A. Dehareng, F., Soyeurt, H. Gillon, A. & Gengler, N. . (2025) Advancing genomic evaluation for Methane Efficiency in Walloon Holstein cattle towards implementation. *Interbull Bull, 61, 182-188.*

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., Vanderick S., Hammami H., Mota R.R., Gillon, A., Hubin, X., Bertozzi, C. and Gengler, N. 2018. Bayesian single-step genomic evaluations combining local and foreign information in Walloon Holsteins. *Animal* 12(5):898–905.

Croquet, C., Mayeres, P., Gillon, A., Vanderick, S., and Gengler, N. 2006. Inbreeding Depression for Global and Partial Economic Indexes, Production, Type, and Functional Traits. *J. Dairy Sci.* 89:2257–2267.

Gao, H., A.A. Kudinov, M. Taskinen, T.J. Pitkänen, M.H. Lidauer, E.A. Mäntysaari, and I. Strandén. 2023. A computationally efficient method for approximating reliabilities in large-scale single-step genomic prediction. *Genet. Sel. Evol*. 55:1.

Garrick, D.J., Taylor, J.F. & Fernando R.L. 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses *Genet. Sel. Evol* 41, 55

Gengler, N., and Vanderick, S. 2008. Bayesian inclusion of external evaluations into a national evaluation system: application to milk production traits. *Interbull Bull., 38. pp. 70-74*

Guarini, A.R., Lourenco, D.A.L., Brito, L.F., Sargolzaei, M.; Baes, C.F., Miglior, F, Tsuruta, S, Misztal, I, Schenkel, F.S. 2019. Use of a single-step approach for integrating foreign information into national genomic evaluation in Holstein cattle, *J. Dairy Sci.* 102: 8175-8183,

Himmelbauer, J, Schwarzenbacher, H., Fuerst, C. 2021. Implementation of single-step evaluations for fitness traits in the German and Austrian Fleckvieh and Brown Swiss populations. *Interbull Bull.*, 56, 82-89.

Lourenco DA, Tsuruta S, Fragomeni BO, Masuda Y, Aguilar I, Legarra A, Bertrand JK, Amen TS, Wang L, Moser DW, Misztal I. 2015. Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. *J Anim Sci.*, 93(6):2653-62.

Mäntysaari, E.A., Koivula, M. , Strandén, I 2020. Symposium review: Single-step genomic evaluations in dairy cattle, *J. Dairy Sci.,* 103, 5314-5326.

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

Misztal, I., S. Tsuruta, D.A.L. Lourenco, I. Aguilar, A. Legarra, and Z.V. 2014. Manual for BLUPF90 family of programs.

Mrode, R.A and Swanson, G.T.J. 1999. Simplified equations for evaluations of bulls in the Interbull international evaluation system. *Livestock Production Science*. 61: 43-52.

Pimentel, E.D.C.G., Edel, C., Krogmeier, D., Emmerling, R. and Götz K.U. 2021. Effects of use of external information in Single-Step evaluations for linear type traits in Brown Swiss. *Interbull Bull.*, 56, 121-124.

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

Stranden, I., M. Lidauer, E. A. Mäntysaari, and J. Pösö. 2000. Calculation of Interbull weighting factors for the Finnish test day model. *Interbull Bull.* 26, 78–79.

Vandenplas, J., Colinet, FG, Glorieux, G., Bertozzi, C., and Gengler, N. 2015. Integration of external estimated breeding values and associated reliabilities using correlations among traits and effects. *J. Dairy Sci.* , 98, 9044-9050.

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

Vandenplas, J., Spehar, M., Potocnik, K., Gengler, N., and Gorjanc, G. 2017. National single-step genomic methods that integrates multi-national genomic information. *J. Dairy Sci.* ,100: 465-78.

Wiggans, G.R. Sonstegard, T.S. VanRaden, P.M., Matukumalli, L.K. Schnabel, R.D., Taylor, J.F., Schenkel, F.S. Van Tassell, C.P. 2009. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the United States and Canada. *J. Dairy Sci.* 92 (7), 3431-3436.

Zaabza, H.B., M. Taskinen, E.A. Mäntysaari, T. Pitkänen, G.P. Aamand, and I. Strandén. 2022. Breeding value reliabilities for multiple-trait single-step genomic best linear unbiased predictor. *J. Dairy Sci.* 105:5221–5237.

Zavadilová, L., Přibyl, J. Vostrý, L. Bauer, J. 2014. Single-step genomic evaluation for linear type traits of Holstein cows in Czech Republic. *Anim. Sci. Pap. Rep.,* 32, 201-208