

Development of a single-step genomic evaluation for udder support in a multi-breed population with crossbred animals

Adrien M. Butty, Peter von Rohr, Patrick Stratz, Urs Schuler, Madeleine Berweger, Urs Schnyder, and Mirjam Spengeler*

Qualitas AG, Chamerstrasse 56, 6300 Zug, Switzerland

**Corresponding author: adrien.buty@qualitasag.ch*

Abstract

In Swiss dairy cattle, pedigree-based breeding values are predicted for Holstein (HO), Swiss Fleckvieh (SF) and Simmental (SI) in a combined evaluation. Although Holstein and Simmental have no recent common genetic history this combined evaluation is possible and stable due to the link established by SF, a cross between HO and SI. The multi-breed evaluation allows higher EBV accuracies particularly for SI as this population is relatively small.

Development of a single-step evaluation for dairy cattle in Switzerland is currently in progress. Udder support is used as example trait. The current multi-breed and crossbred setup must be kept leveraging the benefits of improved accuracies of the combined evaluation. Multiple methods were presented to tackle the problems imposed by a single-step multi-breed evaluation. Using MiX99, a ssGTaBLUP approach was used with all phenotypic and genotypic data available for HO, SF, and SI animals. A residual polygenic term was included at 10% as well as genetic groups specific to each breed. All genotypes were imputed to a density of 125K SNP and reliabilities were estimated according to Ben Zaabza et al. (2020). Scenarios were compared that included phenotypes and genotypes of only one breed or of different combinations of breeds. The aim of the different scenarios was to determine whether the genetic link provided by the crossbred animals could be correctly accounted for in our multi-breed evaluation. Results showed that single-step multi-breed evaluation led to less overestimation of candidate bulls EBV and lower EBV biases in all cases but for the exception of HO candidates. Further work will focus on this breed to improve its validation results.

Key words: Single-step, multi-breed, udder support, crossbred

Introduction

Development of single-step methods for routine genetic evaluations in dairy cattle is in progress in many countries. Implementation of a single-step method is expected to lead to higher accuracies of the estimated breeding values (EBV) as information of the cow genotypes is fully included in the evaluation but also to tackle the overestimation of EBV of young bull candidates observed currently as genomic preselection can be accounted for (Jibrila et al., 2020). In Swiss dairy cattle, pedigree-based breeding values are estimated

for Holstein (HO), Swiss Fleckvieh (SF) and Simmental (SI) in a combined evaluation. In Switzerland, SI is considered a dual-purpose breed for milk and meat and can be compared with, for instance, the German Fleckvieh. Although Holstein and Simmental have no recent common genetic history this combined evaluation is possible and stable due to the link established by SF, a cross between HO and SI. The multi-breed evaluation allows higher accuracies particularly for SI as this population is relatively small.

Multiple methods were presented to tackle the problem imposed by a single-step multi-

breed evaluation: modeling of genetic groups and metafounders are two examples of this (Masuda et al., 2022).

In this study, pedigree BLUP (PBLUP) was compared with ssGTaBLUP (Mäntysaari et al., 2017) in single-breed and in multi-breed setups. The aim of these comparisons was to find out whether the current multi-breed traditional evaluation can be kept once a single-step method is implemented.

Materials and Methods

Data and model

Phenotypic data for the trait udder support were retrieved from the April 2021 routine evaluation and comprised, after quality controls, 1'080'613 data points recorded between January 1992 and February 2021. Genotypes of 364'454 animals were included in the single-step evaluations (Table 1). Phenotypic data, although consisting of discrete scores from 1 (weak) to 9 (strong) with an optimum at 5 can be approximated with a normal distribution and was therefore evaluated with a linear model (Figure 1). As animals were genotyped with different SNP panels, all genotypes were imputed together (one reference panel) to 125K SNP following the routine imputation process at Qualitas with FImpute (v2.2; Sargolzaei et al., 2014). A pedigree was constructed from the phenotyped and genotyped animals up to two generations and comprised 2'114'573 individuals.

Table 1. Percentage of the phenotypes and genotypes across the three breeds

	HO	SF	SI
Phenotypes	79%	18%	3%
Genotypes	94%	2%	4%
Both	84%	10%	6%

Population structure was assessed with a principal component analysis (PCA) with snp1101 (v1.0; Sargolzaei, 2014) on all genotypes.

The described dataset was analyzed with an animal model including classifier-year, year-season, age at calving, lactation stage, heterosis, recombination loss and description scheme as fixed effects and herd, animal and residual as random effects. All evaluations were run with MiX99 (v22.0228; Strandén and Lidauer, 1999) and genetic groups were integrated following the Quaas—Pollak full transformation (Quaas and Pollak, 1981) not only into the pedigree relationship matrix but also in the merged part of genomic and pedigree matrix following Matilainen et al. (2018). Genetic groups were divided by breed but also separated over different periods of time and selection paths. The residual polygenic variance term for the single step evaluation was set to 10%.

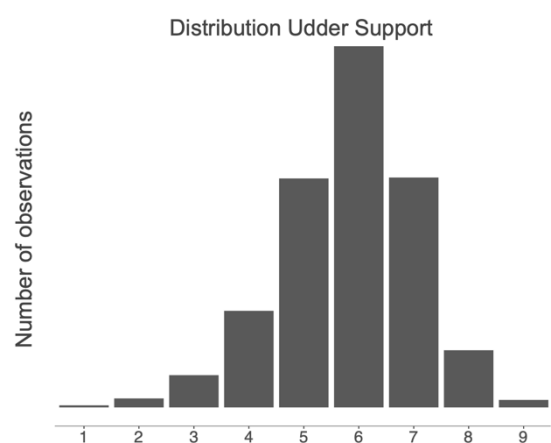


Figure 1. Distribution of USU on a score scale from 1 to 9 with 5 as optimum.

Scenarios and validation

Four different evaluations were run, one with all data at once (multi-breed) and three single-step evaluations for HO, SF, and SI separately.

The LR Method (Legarra and Reverter, 2018) was used to compare the scenarios. Bias and dispersion were analysed for each scenario. The bias is the level difference between the EBV from reduced and the full dataset. In order to facilitate comprehension, the biases were expressed in unit of EBV standard deviation. The dispersion is the regression slope of the EBV of the whole on the partial dataset. For this, data of the four last years were removed

and breeding values estimated with full and reduced datasets were compared for the same validation bulls. Expected values were 0 for the bias and 1 for the dispersion. Validation bulls had an effective record contribution of 0 in the reduced and over 10 in the full datasets. Dispersion values below one indicate an overestimation of the EBV at the time an animal receives its first results relying only on genomic information in comparison to its later EBV relying additionally on progeny information.

Results & Discussion

Population structure

PCA showed that although HO and SI animals are genetically different, SF – their cross – had genetic link of various degrees with both populations (Figure 2).

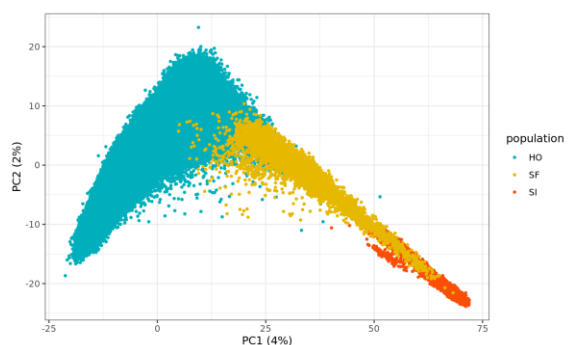


Figure 2. PCA shows how the SF population links HO and SI together.

Traditional and single-step evaluation

Looking at the validation bulls, bias was lower or at the same level for the single-step than for the PBLUP evaluation for all breeds but for the single-step HO scenario (Figure 3A). Dispersion was always better with single-step (Figure 3B). This study shows how implementing single-step method is beneficial as it reduced overestimation of EBV of young bull candidates, even though it does not remove it completely. In both cases, bias was lowest for the multi-breed scenario, which is a first benefit of the current approach in comparison to multiple single-breed models.

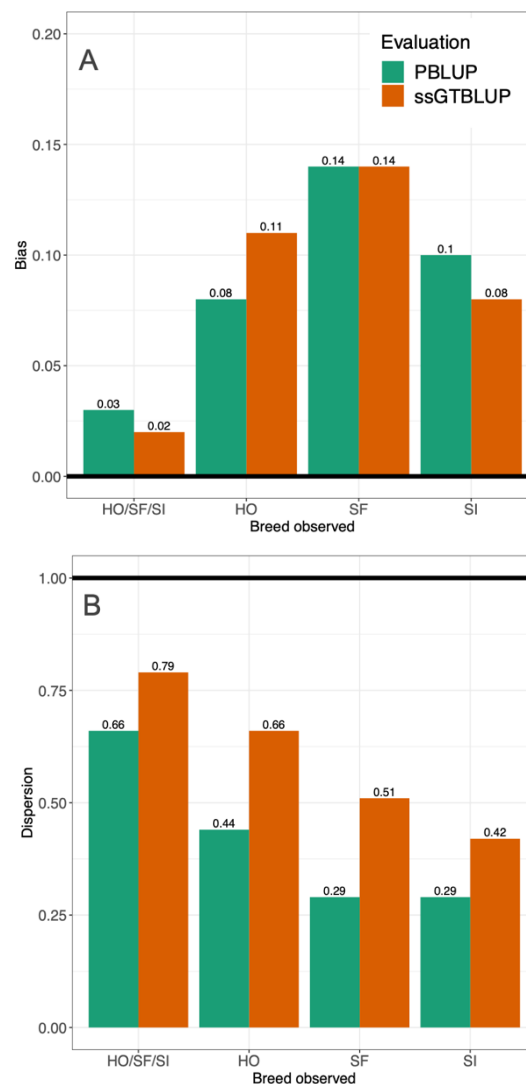


Figure 3. Bias (A) and dispersion (B) observed on the validation bulls for pedigree-based and single-step evaluation

Most of the phenotypes and genotypes in the dataset analyzed in this study were from animals of the breed HO. This is on the one hand due to a larger population in Switzerland but on the other hand also due to the participation of Switzerland to the CDDR genotype exchange program (Weigel, 2014). Through this program, all HO genotypes of males in North-America, Italy and Switzerland are available for Switzerland. This type of exchange does not include SF nor SI animals. The high proportion of young, genotyped HO animals without phenotype and with a low connectedness to the phenotyped Swiss HO population coming from this exchange program

can explain the higher bias when their genotypes are accounted for in the single-step setup.

Experience from other countries showed that animals loosely connected to the phenotyped population had more biased EBV and lower reliabilities (Reiner Emmerling, personal communication). The lack of genomic connectedness and the following diminished EBV quality can possibly be redressed with an increased residual polygenic variance term.

Single- vs multi-breed single-step evaluation

Comparing multi-breed and single-breed single-step evaluations for each population led to the same results as when comparing pedigree and single-step BLUP methods: dispersion values are at least the same or closer to one when all data are accounted for at once (Figure 4B). Bias, in contrary, is only reduced for the SI population when all data are included.

The greater bias observed with single-step multi-breed evaluations for HO and SF can, again, be explained by the specific data structure and origin of most HO genotypes. This could also lead to the greater bias in the SF population as SF animals often have a tighter link to the HO population. The SF population originate from a F1 cross of HO and SI animals but is currently more related to HO than SI on average.

Next steps

Future work should determine the effect of the HO genotypes obtained through the international CDDR data exchange. Removal of genotypes from male candidates not selected for breeding was already envisaged but approaches presented by Koivula et al. (2022) suggested to remove genotype of older animals from the evaluation instead. Both options will be looked at as well as different level of RPG term included in the model.

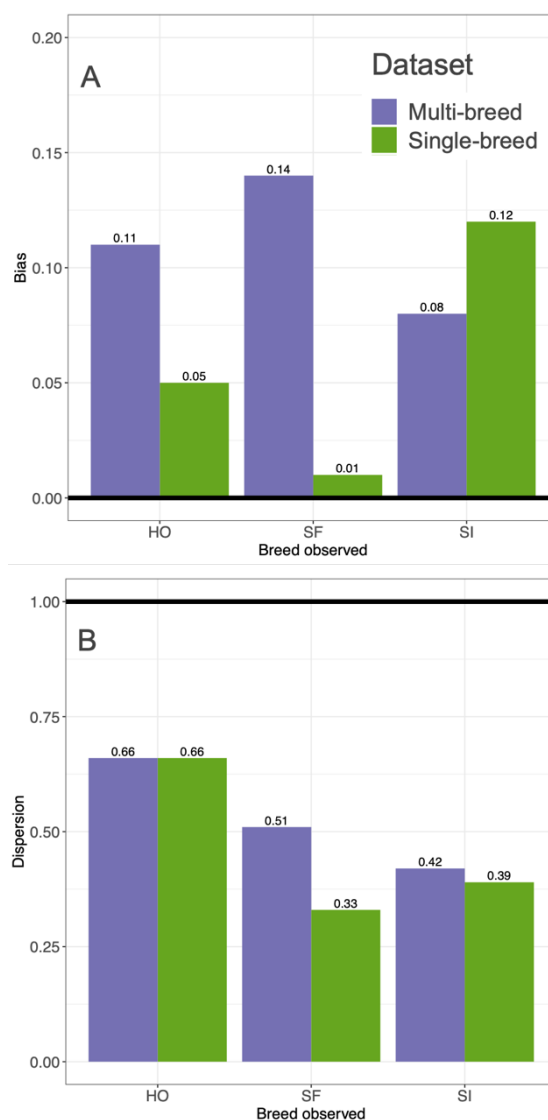


Figure 4. Bias (A) and dispersion (B) observed on the validation bulls for single- and multi-breed single-step evaluations.

Conclusions

Comparison of pedigree and single-step BLUP methods showed improved validation results with the single-step ssGTaBLUP approach. The use of all data at once – multi-breed single-step evaluation – led to less overdispersion of the EBV of young male candidates compared to single-breed datasets. Bias of multi-breed evaluation, however, was greater for HO and SF population. Further work must therefore investigate how to reduce these biases.

Acknowledgments

We greatly thank the MiX99 team of Luke, the Natural Resources Institute of Finland for their support and availability of the latest development MiX99 version. The association of Swiss cattle breeding organisation, and in particular Swissherdbook and Holstein Switzerland, are acknowledged for providing the data used in this study.

References

- Ben Zaabza, H., M. Taskinen, T. Pitkänen, G.P. Aamand, E.A. Mäntysaari, and I. Strandén. 2020. Approximate individual animal reliabilities in single-step model. Page in EAAP Book of Abstract 2020.
- Jibrila, I., J. Napel, J. Vandenplas, R.F. Veerkamp, and M.P.L. Calus. 2020. Investigating the impact of preselection on subsequent single - step genomic BLUP evaluation of preselected animals. *Genet. Sel. Evol.* 52:1–10. doi:10.1186/s12711-020-00562-6.
- Koivula, M., U.S. Nielson, I. Strandén, G.P. Aamand, and E.A. Mäntysaari. 2022. Nordic Holstein single-step test day model using left truncated genomic data. Page in Interbull Meeting.
- Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genet. Sel. Evol.* 50:1–18. doi:10.1186/s12711-018-0426-6.
- Mäntysaari, E.A., R.D. Evans, and I. Strandén. 2017. Efficient single-step genomic evaluation for a multibreed beef cattle population having many genotyped animals. *J. Anim. Sci.* 95:4728–4737. doi:10.2527/jas2017.1912.
- Masuda, Y., P.M. VanRaden, S. Tsuruta, D.A.L. Lourenco, and I. Misztal. 2022. Invited review: Unknown-parent groups and metafounders in single-step genomic BLUP. *J. Dairy Sci.* 105:923–939. doi:10.3168/jds.2021-20293.
- Matilainen, K., I. Strandén, G.P. Aamand, and E.A. Mäntysaari. 2018. Single step genomic evaluation for female fertility in Nordic Red dairy cattle. *J. Anim. Breed. Genet.* 135:337–348. doi:10.1111/jbg.12353.
- Quaas, R.L., and E.J. Pollak. 1981. Modified Equations for Sire Models with Groups. *J. Dairy Sci.* 64:1868–1872. doi:10.3168/jds.S0022-0302(81)82778-6.
- Sargolzaei, M. 2014. SNP1101 User's Guide. Version 1.0.. HiggsGene Solut. Inc.
- Sargolzaei, M., J.P. Chesnais, and F.S. Schenkel. 2014. A new approach for efficient genotype imputation using information from relatives.. *BMC Genomics* 15:478. doi:10.1186/1471-2164-15-478.
- Strandén, I., and M. Lidauer. 1999. Solving large mixed linear models using preconditioned conjugate gradient iteration. *J. Dairy Sci.* 82:2779–2787. doi:10.3168/jds.S0022-0302(99)75535-9.
- Weigel, K.A. 2014. The Cooperative Dairy DNA Repository (CDDR) How it has helped the AI Industry. Page in NAAB Technical Conference.