

# Development of an approximate multivariate two-step approach for the joint genetic evaluation of Austrian and German dairy cattle

**Pfeiffer, C., Fuerst-Waltl, B., Fuerst, C.**

**Poljoprivreda/Agriculture**

ISSN: 1848-8080 (Online)

ISSN: 1330-7142 (Print)

<http://dx.doi.org/10.18047/poljo.21.1.sup.7>



**Poljoprivredni fakultet u Osijeku, Poljoprivredni institut Osijek**

Faculty of Agriculture in Osijek, Agricultural Institute Osijek

# DEVELOPMENT OF AN APPROXIMATE MULTIVARIATE TWO-STEP APPROACH FOR THE JOINT GENETIC EVALUATION OF AUSTRIAN AND GERMAN DAIRY CATTLE

Pfeiffer, C., Fuerst-Waltl, B., Fuerst, C.

Original scientific paper

## SUMMARY

**Multivariate genetic evaluation in modern dairy cattle breeding programs became important in the last decades. The simultaneous estimation of all production and functional traits is still demanding. Different meta-models are used to overcome several constraints. The aim of this study was to conduct an approximate multivariate two-step procedure applied to de-regressed breeding values and yield deviations of five fertility traits of Austrian Pinzgau cattle and to compare results with routinely estimated breeding values. The approximate two-step procedure applied to de-regressed breeding values performed better than the procedure applied to yield deviations. Spearman rank correlations for all animals, sires and cows were between 0.996 and 0.999 for the procedure applied to de-regressed breeding values and between 0.866 and 0.995 for the procedure applied to yield deviations. The results are encouraging to move from the currently used selection index in routine genetic evaluation towards an approximate two-step procedure applied to de-regressed breeding values.**

**Key-words: approximate multiple trait, de-regressed breeding values, yield deviation, fertility, cattle**

## INTRODUCTION

Due to the increased number of production and functional traits in modern dairy cattle breeding programs, multivariate genetic evaluation became increasingly interesting over the last decades. One of the major challenges is the simultaneous evaluation of all traits. Hence, different meta-models were proposed for national and international genetic evaluations, e.g. an approximate two-step approach using pseudo-phenotypes (Ducrocq et al., 2001) or Multiple trait Across Country Evaluation (MACE; Schaeffer, 1994; Schaeffer, 2001). At present, the joint genetic evaluation of Austria and Germany is optimised aiming at a multiple trait genetic evaluation. Currently, selection is based on a total merit index (TMI) derived by Miesenberger (1997). The TMI of the evaluated cattle breeds consists up to 30 production and functional traits.

Breeding values for the TMI as well as for several sub-indices are estimated either using univariate (e.g. protein yield) or multivariate (e.g. calving ease and stillbirth) methods by applying animal or sire-maternal-

grandsire models (the latter for functional longevity only). Some of these models include repeated measures, such as somatic cell count (Fuerst et al., 2015). Subsequently, estimated breeding values (EBV) for individual traits are combined to form the TMI or other sub-indices by assuming that residual covariances between traits or groups of traits are equal to zero. Additionally, genetic correlations between many traits are assumed to be zero or were obtained from literature (Miesenberger, 1997). These constraints lead to an upwards biased TMI for animals with low reliabilities. This assumption was confirmed on simulated data by Pfeiffer et al. (2015). In fact, a full multivariate estimation based on phenotypic data would be the optimum methodology. However, it is usually not feasible (Mrode, 2014) due to the tremendous amount of data in genetic evaluations and restricted computer power (Lassen et

(1) Christina Pfeiffer, Dipl. Ing. (christina.pfeiffer@boku.ac.at), Dr. Birgit Fuerst-Waltl - University of Natural Resources and Life Sciences Vienna (BOKU), Division of Livestock Sciences, Gregor-Mendel-Str. 33, 1180 Vienna, Austria, (2) Dr. Christian Fuerst – ZuchtData EDV-Dienstleistungen GmbH, Dresdner Straße 89/19, 1200 Vienna, Austria

al., 2007). Thus, an approximate multivariate model using a two-step procedure was proposed and validated using simulated data (Durcrocq et al., 2001; Lassen et al., 2007; Pfeiffer et al., 2015). Results of the simulation study by Pfeiffer et al. (2015) were encouraging, but the procedure has to be approved on field data. Therefore, the objective of this study was the comparison of routinely estimated breeding values (EBVr) for fertility traits with results of an approximate two-step procedure based on de-regressed breeding values (drEBV) and yield deviations (YD), respectively.

## MATERIAL AND METHODS

All routinely evaluated fertility traits of Pinzgau cattle were chosen to test the approximate two-step procedure applied to drEBV (procedure PdrEBV) and YD (procedure PYD). These were non-return-rate 56 for heifers (NR-H) and cows (NR-C), days from first calving to first insemination (CFI) and days from first to last insemination for heifers (FLI-H) and cows (FLI-C). In total, 294,027 records of 104,866 cows and heifers inseminated between the years 1990 and 2014 were analysed. The pedigree consisted of 183,129 animals. In the first step for PdrEBV, a 5-trait genetic evaluation was applied to get EBV using the program package MiX99 (Lidauer et al., 2015). The following statistical model was used:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{e} \quad (1)$$

where  $\mathbf{y}$  is a vector of observations of the traits NR-H, NR-C, CFI, FLI-H and FLI-C;  $\mathbf{b}$  is a vector of systematic effects, including fixed effect of region-year-month of insemination interaction, herd-year interaction, parity-age at calving/insemination interaction, inseminator-year (only for NR-H and NR-C) interaction and service sire (only for NR-H and NR-C);  $\mathbf{a}$  is a vector of animal effects;  $\mathbf{p}$  is a vector of permanent environmental effects (only for cows) and  $\mathbf{e}$  is a vector of residuals;  $\mathbf{X}$ ,  $\mathbf{Z}$  and  $\mathbf{W}$  are the corresponding incidence matrices. Fertility EBV were then de-regressed by a multivariate de-regression approach (Schaeffer, 2001), implemented in the program package MiX99 (Lidauer et al., 2015). The de-regression procedure uses the estimated breeding values and their respective effective daughter contributions as weights only considering the general mean as fixed effect. For the second procedure PYD, YD were computed, again using the software MiX99 (Lidauer et al., 2015). The following model was applied:

$$\mathbf{y}^* = \mathbf{y} - \mathbf{Xb} - \mathbf{Zp} \quad (2)$$

where  $\mathbf{y}^*$  is a vector of YD;  $\mathbf{y}$  is a vector of phenotypic observations of the traits NR-H, NR-C, CFI, FLI-H and FLI-C;  $\mathbf{b}$  indicates the vector of all fixed effects, already described for equation 1;  $\mathbf{p}$  is a vector of permanent

environmental effects (only for cows);  $\mathbf{X}$  and  $\mathbf{Z}$  are the corresponding incidence matrices.

After de-regressing EBV and computing YD, respectively, all five traits were analysed by means of the following multivariate animal model:

$$\mathbf{y}^\# = \boldsymbol{\mu} + \mathbf{Za} + \mathbf{e} \quad (3)$$

where  $\mathbf{y}^\#$  indicates either drEBV or YD of the respective trait;  $\boldsymbol{\mu}$  is the general mean;  $\mathbf{a}$  is a vector of random animal additive genetic effects and  $\mathbf{e}$  denotes a vector of random residual effects.  $\mathbf{X}$  and  $\mathbf{Z}$  represent the corresponding incidence matrices. Based on approximate Interbull reliabilities (Strandén et al., 2000) effective own performances (Edel et al., 2009) were calculated and used as weighting factors for drEBV and YD in the multivariate genetic evaluation. Routine genetic parameters were used (Fuerst et al., 2015) for the multivariate genetic evaluation. Spearman rank correlations between EBV of routine genetic evaluation, PdrEBV and PYD were calculated using the program package SAS 9.2 (SAS, 2008). All EBV were standardised to relative breeding values with a mean of 100 and an additive genetic standard deviation of 12.

## RESULTS AND DISCUSSION

Means and standard deviations of EBVr, PdrEBV and PYD for all animals and sires with reliabilities higher than 50% ( $n=318$ ) are given in Table 1. Means for each trait and procedure were similar, also standard deviations of EBVr and PdrEBV were almost equal. Standard deviations for PYD were lower compared to EBVr and PdrEBV. Table 2 shows the rank correlations between routinely estimated breeding values and the two-step procedure applied to drEBV and YD, respectively, for all animals, sires with reliabilities higher than 50% and cows with reliabilities higher than 30%. Correlations between EBVr and PdrEBV were almost 1 for all traits and animal groups. Correlations between EBVr and PYD were lower. These results were in accordance with the simulation study of Pfeiffer et al. (2015). Authors could show that outcomes of an approximate two-step procedure applied to drEBV were always closer to the reference method, which was a full multivariate animal model based on phenotypic data, than those of an approximate two-step procedure applied to YD.

**Table 1. Means and standard deviations of the routinely estimated breeding values, breeding values derived from an approximate two-step procedure based on de-regressed breeding values and yield deviations for non-return-rate 56 heifer (NR-H) and cow (NR-C), days from first calving to first insemination (CFI) and days from first to last insemination for heifers (FLI-H) and cows (FLI-C)**

	Estimated breeding values				De-regressed breeding values				Yield deviations			
	$x_a$	$s_a$	$x_s$	$s_s$	$x_a$	$s_a$	$x_s$	$s_s$	$x_a$	$s_a$	$x_s$	$s_s$
NR-H	99.4	4.2	99.8	6.2	100.0	4.1	99.8	6.2	99.7	3.0	99.7	5.7
NR-C	99.0	4.0	99.4	6.6	99.0	4.0	99.4	6.5	99.0	3.5	99.3	6.3
CFI	104.8	6.7	102.5	11.5	104.8	6.6	102.5	11.4	104.5	6.1	102.4	10.8
FLI-H	100.7	5.0	100.5	7.2	100.7	4.9	100.5	7.1	100.7	3.8	100.4	6.9
FLI-C	101.0	5.5	100.1	9.0	101.0	5.4	100.1	8.9	100.7	4.9	99.9	8.6

$x_a$  = mean of all animals,  $s_a$  = standard deviation of all animals;  $x_s$  = means of sires with reliabilities >50%;  $s_s$  = standard deviation of sires with reliabilities >50%

**Table 2. Rank correlations between routinely estimated breeding values (EBVr), breeding values derived from an approximate two-step procedure based on de-regressed breeding values and yield deviations for non-return-rate 56 heifer (NR-H) and cow (NR-C), days from first calving to first insemination (CFI) and days from first to last insemination for heifers (FLI-H) and cows (FLI-C) for all animals (N=183,129), sires with reliabilities >50% (N=318) and cows with reliabilities >30% (N=12,792)**

Trait	De-regressed breeding values			Yield deviations		
	All	Sires	Cows	All	Sires	Cows
NR-H	0.997	0.999	0.998	0.866	0.967	0.970
NR-C	0.996	0.999	0.998	0.966	0.989	0.984
CFI	0.998	0.999	0.999	0.983	0.995	0.988
FLI-H	0.998	0.999	0.999	0.937	0.987	0.977
FLI-C	0.996	0.999	0.998	0.982	0.992	0.987

In accordance with the earlier studies (Sigurdsson and Banos, 1995; Thomsen et al., 2001), proposing drEBV to be reliable alternatives to daughter yield deviations, the approximate two-step procedure applied to drEBV is feasible. For the development of routine genetic evaluation, an approximate two-step procedure applied to drEBV is recommended as drEBV are easier available for all traits included in the TMI than YD and Interbull breeding values can be implemented straightforwardly. Unpublished results of Pfeiffer (2015) also showed that the estimation of genetic parameters using an approximate two-step procedure applied to drEBV was feasible. The entire procedure, including new genetic parameters is still under development.

## CONCLUSION

An approximate two-step procedure applied to drEBV and YD based on field data is feasible. The results are encouraging for further work on their implementation in routine genetic evaluation. The results open up perspectives for the replacement of the current selection index method by an approximate two-step procedure based on drEBV.

## ACKNOWLEDGEMENT

Authors greatly acknowledge the funding of the project OptiGene ('Optimization of long-term genetic progress of Austrian cattle breeds with emphasis on health and genomic selection'; project number 100808) by the Austrian Federal Ministry of Agriculture, Forestry, Environment and Water Management, the Federations of Austrian Fleckvieh, Brown-Swiss, Pinzgauer and Tyrolean Grey cattle and the Federation of Austrian Cattle Breeders (ZAR).

## REFERENCES

1. Ducrocq, V., Boichard, D., Barbat, A., Larroque, H. (2001): Implementation of an approximate multitrait BLUP evaluation to combine production traits and functional traits into a total merit index. Proc. 52nd Annual Meeting EAAP, Budapest, Hungary.
2. Edel, C., Emmerling, R., Götz, K-U. (2009): Optimized aggregation for MA-BLUP evaluation in German Fleckvieh. Proc. of the Interbull Meeting, Barcelona, Spain; Interbull Bulletin 40: 178-83.
3. Fuerst, C., Dodenhoff, J., Egger-Danner, C., Emmerling, R., Hamann, H., Krogmeier, D., Schwarzenbacher, H. (2015): Zuchtwertschätzung beim Rind - Grundlagen,

- Methoden und Interpretation. <http://www.zar.at/download/ZWS/ZWS.pdf>. Accessed 15 April 2015.
4. Lassen, J., Sørensen, M.K., Madsen, P., Ducrocq, V. (2007): An approximate multitrait model for genetic evaluation in dairy cattle with a robust estimation of genetic trends. *Genetic Selection Evolution* 39: 353-67. doi: <http://dx.doi.org/10.1051/gse:2007008>
  5. Lidauer, M., Matilainen, K., Mäntysaari, E., Pitkänen, T., Taskinen, M., Strandén, I. (2015): MiX99 – Solving large mixed model equations. Release 1/2015, Jokioinen, Finland.
  6. Miesenberger, J. (1997): Zuchtzieldefinition und Indexselektion für die österreichische Rinderzucht. PhD thesis. University of Natural Resources and Life Sciences Vienna, Austria.
  7. Mrode, R.A. (2014): Linear models for the prediction of animal breeding values. 3rd edition. CABI. Wallingford, UK.
  8. Pfeiffer, C., Fuerst-Waltl, B., Schwarzenbacher, H., Steininger, F., Fuerst, C. (2015): A comparison of methods to calculate a total merit index using stochastic simulation. *Genetic Selection Evolution* 47: 36. doi: <http://dx.doi.org/10.1186/s12711-015-0118-4>
  9. SAS (2008). User's Guide. SAS Institute Inc., Cary, NC, USA.
  10. Schaeffer, L.R. (1994): Multiple-country comparison of dairy sires. *Journal of Dairy Science*, 77: 2671-2678. doi: [http://dx.doi.org/10.3168/jds.S0022-0302\(94\)77209-X](http://dx.doi.org/10.3168/jds.S0022-0302(94)77209-X)
  11. Schaeffer, L.R. (2001): Multiple trait international bull comparison. *Livestock Production Science*, 69: 145-153. doi: [http://dx.doi.org/10.1016/S0301-6226\(00\)00255-4](http://dx.doi.org/10.1016/S0301-6226(00)00255-4)
  12. Sigurdsson, A., Banos, G. (1995): Dependent variables in international sire evaluation. *Acta Agriculturae Scandinavica*, 45: 209-217. doi: <http://dx.doi.org/10.1080/09064709509413079>
  13. Strandén, I., Lidauer, M., Mäntysaari, E.A., Pösö, J. (2000): Calculation of Interbull weighting factors for the Finnish test day model. Proc. of the Interbull technical workshop, Verden, Germany; *Interbull Bulletin* 26: 78-79.
  14. Thomsen, H., Reinsch, N., Xu, N., Looft, C., Grupe, S., Kühn, C. et al. (2001): Comparison of estimated breeding values, daughter yield deviations and de-regressed proofs within a whole genome scan for QTL. *Journal of Animal Breeding and Genetics*, 118: 357-370. doi: <http://dx.doi.org/10.1046/j.1439-0388.2001.00302.x>

(Received on 27 April 2015; accepted on 6 July 2015)