

Comparison of different methods to calculate a total merit index - results of a simulation study

C. Fuerst¹, C. Pfeiffer², H. Schwarzenbacher¹, F. Steininger¹, B. Fuerst-Waltl²

¹ ZuchtData, Vienna, Austria, ² University of Natural Resources and Life Sciences, Vienna, Austria

ABSTRACT: A stochastic simulation study was conducted to analyze four different methods to combine five traits into a total merit index. The reference method was a multitrait evaluation based on raw data. Two methods were variations of an approximate multitrait model; one based on yield deviations, the other on de-regressed estimated breeding values (EBVs). The fourth method was an adaptation of the selection index that is used in the routine evaluation in Austria and Germany. Both approximate multitrait models turned out to be very close to the reference model. The selection index method showed good results in most cases but led to a noticeable bias in EBVs in the case of high residual covariances. The results of this study are encouraging to move towards the approximate multitrait approach, but further analyses are necessary before it can be implemented in the routine evaluation.

Keywords: stochastic simulation; total merit index; genetic evaluation; multitrait model; selection index

Introduction

Modern dairy cattle breeding goals include several production and increasingly more functional traits. Estimated breeding values (EBVs) of the traits that are combined in the total merit index (TMI), usually come from single trait models or from multivariate models for groups of traits (e.g. Fuerst et al., 2013). It is well known that the reliability of the EBV can be increased by using multitrait models to fully exploit information of correlated traits (e.g. Thompson and Meyer, 1986). In most cases, a multivariate animal model based on raw data for all traits in the total merit index is not feasible. Therefore, proper approximations are needed in the routine evaluation. Ducrocq et al. (2001) and Lassen et al. (2007) proposed an approximate multitrait model using a two-step procedure. The first step is a single trait model for all traits to calculate yield deviations (YDs) for each animal. In the second step a multitrait animal model, correcting for the random genetic animal and a fixed year effect, is applied.

In the joint genetic evaluation in Austria and Germany, a method based on selection index theory (Hazel and Lush, 1942) proposed by Miesenberger (1997) is in use for the TMI and several sub-indices for all breeds except Holstein.

The aims of this study were (1) to simulate data for five traits and a simplified but representative cattle population scheme and (2) to compare different methods to calculate a TMI.

Materials and Methods

Simulation. A population structure roughly modeling Austrian Brown Swiss cattle was simulated with

the stochastic simulation program ADAM (Pedersen et al., 2009). The simulated population size was approximately 50,000 cows distributed on 3,420 herds. Five normally distributed traits were chosen to represent the trait blocks milk, beef and fitness, with a wide range of heritabilities and genetic correlations. Four traits, fat yield, protein yield, somatic cell count and non-return rate, were measured on all females and net daily gain on approximately 85% of all male animals. No repeated records were assumed. A selection scheme with 25% young bulls and 75% proven bulls was simulated. Breeding values and phenotypes for the five traits were simulated for base population animals. Afterwards, animals were selected on a TMI based on multivariately estimated EBVs over 20 years. Relative economic weights per genetic standard deviation for fat yield, protein yield, net daily gain, somatic cell count and non-return rate were close to the values in the routine evaluation, which are 5.4, 53.6, 4.3, 17.0 and 19.7%, respectively. Heritabilities and genetic correlations between the five traits are shown in Table 1. Three scenarios with respect to the covariances of the residual effects were simulated. In scenarios 0, 1 and 2, residual correlations were varied from zero, to half of, and equal to the genetic correlations, respectively. These scenarios were chosen to specifically test the method of Miesenberger (1997) for a possible bias, because it ignores correlations between residual effects. Three replicates were simulated for each scenario.

Table 1. True genetic parameters (heritabilities on diagonal, genetic correlations above diagonal).

Trait	Fkg	Pkg	NDG	SCC	NRR
Fkg	0.40	0.85	0.10	0.25	-0.20
Pkg		0.39	0.10	0.25	-0.20
NDG			0.27	0.00	0.00
SCC				0.12	-0.10
NRR					0.02

Fkg=fat yield, Pkg=Protein yield, NDG=net daily gain, SCC=somatic cell count, NRR=non return rate

TMI methods. The reference method was a multitrait animal model based on raw data using the true genetic and phenotypic parameters (method MULTI). The statistical model included a fixed herd-year-effect, a random genetic and a random residual effect. Method YD was the approximate multitrait approach proposed by Ducrocq et al. (2001), which is based on YDs. A modification of this approach was applied by using de-regressed EBVs (drEBVs) instead of YDs (method DR). For both methods, YD and DR, univariate genetic evaluations were necessary for each trait to calculate YDs and drEBVs. This was done with the program package

MiX99 (Lidauer et al., 2013). Based on reliabilities using the program ApaX (Stranden et al., 2001), effective own performances were used as weights for YDs and drEBVs. The TMI was calculated by multiplying the multivariately estimated EBVs with the economic weights for all three methods, MULTI, YD and DR. The fourth method was the approach currently used in the routine evaluation. EBVs of the five traits were estimated univariately and then combined by applying the selection index method of Miesenberger (1997). In this adaptation of the selection index method (Hazel and Lush, 1942), EBVs are combined instead of phenotypes (method M). The covariances between the EBVs (σ_{xy}) are calculated as:

$$\sigma_{xy} = r_{gxy} r_x^2 r_y^2 \sigma_{ax} \sigma_{ay}$$

where r_{gxy} = genetic correlation between traits x and y, r_x^2 , r_y^2 = reliabilities of EBVs for traits x and y, and σ_{ax} , σ_{ay} = additive genetic standard deviations of traits x and y, respectively. This means that residual correlations are neglected, i.e. assumed to be zero.

For all methods, genetic parameters were not re-estimated, but the true (simulated) parameters were used. All EBVs were transformed to relative breeding values (RBV) by standardizing to 12 points per additive genetic standard deviation; the base was set to 100 for the years 10 to 15 (RBV 100/12).

Results and Discussion

Correlations. Table 2 shows the Spearman rank correlations between the different TMIs and the true TMI for scenarios 0 and 2. Across all animals from year 1 to 20, the correlations were above 0.93 for all methods that assumed zero residual correlations. These correlations are very high because of the strong genetic trend in this simulation. Therefore the correlations in Table 2 are within year groups to reduce the effect of the genetic trend. Correlations were very similar for the methods MULTI, YD and DR. However, for method M, correlations were lower for all year groups. Overall, rank correlations for scenario 2, where the residual (and thus the phenotypic) correlation was equal to the genetic correlation, were slightly lower for all methods. Results for scenario 1 are not shown, but were in general between scenarios 0 and 2.

Table 2. Rank correlations with true TMI within year groups for different TMI methods for scenarios 0 and 2.

Scen.	Years	MULTI	YD	DR	M
0	All	0.9357	0.9357	0.9357	0.9311
	1-5	0.6395	0.6389	0.6394	0.6168
	6-10	0.6902	0.6900	0.6901	0.6706
	11-15	0.6731	0.6730	0.6730	0.6401
	16-20	0.6344	0.6345	0.6345	0.6050
2	All	0.9238	0.9238	0.9238	0.9174
	1-5	0.6840	0.6835	0.6839	0.6580
	6-10	0.6400	0.6398	0.6400	0.6078
	11-15	0.6281	0.6278	0.6281	0.5942
	16-20	0.6251	0.6247	0.6251	0.5838

Correlations with reference method MULTI for scenario 2 (Table 3) were close to 1 for methods YD and DR, whereas correlations for method M were between 0.93 and 0.96. For scenarios 0 and 1, the correlations were slightly higher for method M than for scenario 2.

Table 3. Rank correlations with multivariate TMI within year groups for different TMI methods for scenario 2.

Years	YD	DR	M
All	0.9999	1.0000	0.9928
1-5	0.9992	0.9999	0.9574
6-10	0.9993	0.9999	0.9479
11-15	0.9993	0.9999	0.9447
16-20	0.9991	0.9999	0.9355

Bias. The bias of the different TMIs from the true TMI was calculated by subtracting the true TMI from the estimated TMI for all animals, both expressed as RBV (100/12). Results for scenarios 0 and 2 are shown in Table 4. With no residual covariance (scenario 0) all methods showed rather good results with respect to bias. Only in the very first years of selection some bias was observed, which is probably due to incomplete pedigrees and the use of phantom parent groups. This was also valid for scenario 2, except for method M. Method M led to a markedly overestimated genetic trend, which is expressed as a downwards bias in the first years and an upwards bias in the last years (Figure 1). This result was even stronger for the best 10% of the animals for TMI per year (Table 5), showing an overestimation of the top animals with method M. This is the result of an inflated variance of the TMI with method M, particularly in the presence of residual covariances.

Table 4. Bias (EBV-TBV) of different TMI methods from true TMI within year groups for scenarios 0 and 2.

Scen.	Years	MULTI	YD	DR	M
0	All	0.4	0.6	0.4	-0.6
	1-5	1.6	2.7	1.9	-1.7
	6-10	0.6	1.1	0.7	-1.3
	11-15	-0.1	-0.1	-0.1	0.0
	16-20	-0.6	-1.2	-0.8	0.5
2	All	0.3	0.6	0.4	-0.9
	1-5	1.0	2.1	1.4	-4.7
	6-10	0.2	0.7	0.4	-2.9
	11-15	0.0	0.0	0.0	0.4
	16-20	0.1	-0.4	-0.1	3.5

Conclusions

Results show that all analyzed methods to calculate a TMI lead to quite high correlations with the true TMI. This is particularly true when residual covariances are zero or low. However, in real data, residual covariances can be relevant. The approximate multitrait approach proposed by Ducrocq et al. (2001) gives results that are very close to multitrait evaluation based on raw data. This is not only valid with YDs, using drEBVs even gives slightly better results. From the results of this simulation, drEBVs could

be a good alternative to YDs, as they might be easier to obtain in some cases (e.g. persistency). This could also help to include Interbull EBVs in national evaluations, as individual YDs are not available at the international level. The currently used method M shows good results for higher reliabilities but leads to inflated deviations with low reliabilities. This results in a bias, particularly for the top animals and is therefore relevant in terms of selection accuracy.

Figure 1. Time trend of bias (EBV-TBV) of different TMI methods for scenario 2.

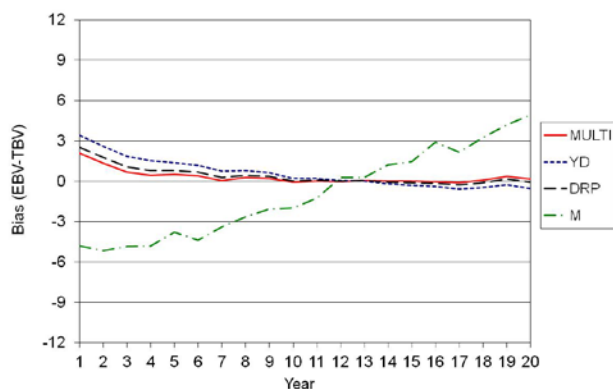


Table 5. Bias (EBV-TBV) of different TMI methods from true TMI for the top 10% within year groups for scenario2

Years	MULTI	YD	DR	M
All	0.3	-0.4	0.1	4.4
1-5	0.8	1.6	1.1	-3.7
6-10	0.0	0.7	0.3	-4.6
11-15	0.5	-0.4	0.2	2.8
16-20	0.3	-0.5	0.0	5.3

index calculation by the approximate multitrait approach. Both versions with YDs and with drEBVs appear to work fine. The crucial point for implementation in routine evaluation is that it is a very difficult task to estimate all genetic and residual (co)variances for all traits in the TMI. For example, 15 different traits with their economic weights plus several indicator traits are currently included in the official TMI for Brown Swiss. Another important point is to incorporate genomic information from the national and international genomic evaluations (e.g. Intergenomics). Therefore, further scenarios and analyses are necessary before implementation into the routine evaluation is feasible.

Acknowledgements

We particularly thank Anders Christian Sørensen and his colleagues from the Aarhus University for granting access and support with the simulation program ADAM. We gratefully 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 no. 100808) by the Austrian Federal Ministry of Agriculture, Forestry, Environment and Water Management and the Federation of Austrian Cattle Breeders (ZAR).

Literature Cited

- Ducrocq, V., Boichard, D., Barbat, A. et al. (2001). Proc. 52nd EAAP meeting, Budapest:3.
- Fuerst, C., Dodenhoff, J., Emmerling, R. et al. (2013). <http://www.zar.at/download/ZWS/ZWS.pdf>.
- Hazel, L.N., Lush, J.L. (1942). *J. Heredity* 33:393-399.
- Lassen, J., Sorensen, M.K., Madsen, P., Ducrocq, V. (2007). *Genet. Sel. Evol.* 39:353-367.
- Lidauer, M., Matilainen, K., Mäntysaari, E., et al. (2013). *MiX99 – Manual*.
- Miesenberger, J. (1997). PhD Thesis, Univ. Agr. Sci. Vienna, Austria.
- Pedersen, L. D., Sørensen, A. C., Henryon, M. et al. (2009). *Livest. Sci.* 121:343-344.
- Stranden, I., Lidauer, M., Mäntysaari, E.A., et al. (2001). *Interbull Bulletin* 26:78-79.
- Thompson, R., Meyer, K. (1986). *Livest. Prod. Sci.* 15:299-313.

The results of this simulation study are encouraging to replace the current method of TMI and sub-