Conversion Equations Revisited:
Single- and Multiple-Trait Conversion Equations and Comparison with Simple-MACE

Hossein Jorjani[1] & Freddy Fikse[1]

[1]*Interbull Centre, Department of Animal Breeding & Genetics, Swedish University of Agricultural Sciences, Box 7023, S - 750 07, Uppsala, Sweden*

Abstract

Results of national genetic evaluation from the USA for 29 traits were used for estimation of converted breeding values for 61 Swedish traits, including the Swedish Total Merit Index (TMI). The Swedish TMI had high correlations with the USA composite traits Cheese Merit $ and Net Merit $ (0.73 and 0.72, respectively). After adjustment for reliabilities these correlations increased to 0.78 and 0.77 respectively. Multiple regression conversion equations had higher predictive ability and lower variance of error of prediction. Different alternatives of Simple-MACE methodology performed better than conversion equations when correlation between the two traits was low. However, Simple-MACE methods performed poorer than conversion equations when used for traits with high correlation or in multiple trait situations.

Introduction

Total merit index (TMI) is used in 25 out of 31 Interbull member countries. The TMI used in Sweden for Holstein bulls comprises 11 different sub-indices and 36 traits, for 18 of which there is no international genetic evaluation (IGE). In the USA different combinations of 11 traits contribute to three kinds of TMI, namely Net Merit, Cheese Merit and Fluid Merit, all of which are expressed in Dollars. Out of these 11 traits there are no IGE for four of them.

Adoption of conversion equations by the International Dairy Federation (1981) and later modifications to the basic concept by Goddard (1985), Wilmink *et al.* (1986) and Powell (1988) took place in a time and environment where national genetic evaluations were restricted to only a couple of production traits. Nowadays, some countries have evaluations for more than 50 traits. Availability of many more traits than before raises the question of whether conversion equations work equally well on traits with low correlation, for cross-trait conversion, or composite traits as they did on yield production traits and whether the method can be extended to more than one trait.

The aims of this study were a) To assess the use of conversion equations for traits with low correlation, for cross-trait predictions, and for composite traits; b) To extend the method of single trait conversion to multiple traits; and c) to compare the effectiveness of single and multiple trait conversions with that of Simple-MACE.

Material and Methods

The data used in this study were results of national genetic evaluations for Holstein dairy bulls from the USA and Sweden. From Sweden relative breeding values (RBV) and from the USA predicted transmitting abilities (PTA) were reported.

Table 1. Number of records.

	All bulls	Bulls born > 1984
Number of bulls	12333	4526
Swedish traits		
Milk yield (max.)	3401	1714
Stature (min.)	520	507
USA traits		
Net Merit $ (max.)	902	539
OFL (min.)	596	431

Number of the so called common bulls, i.e. bulls with one evaluation in each country, was very variable (Table 1) and ranged from a

maximum of 285 bulls between Swedish trait milk yield and USA trait Net Merit $ to a minimum of 74 bulls between Swedish trait stature and USA trait overall feet and leg score.

Single trait regression analysis – Let a group of bulls to have daughters, and consequently estimated breeding values (EBV), in an exporting (EXP) and an importing (IMP) country. Then, a regression equation can be used to describe the relationship between EBVs in the two countries as:

$$EBV_{IMP} = a + b \, EBV_{EXP} + \varepsilon \qquad [1]$$

where a and b are is the intercept and the slope of the regression equation, respectively. Having established this relationship one can predict a converted breeding value (CBV) in the IMP country for bulls from the EXP country that do not yet have any daughters in the IMP country, i.e.

$$CBV_{IMP} = a + b \, EBV_{EXP} \qquad [2]$$

Wilmink *et al.* (1986) suggested a modification to this basic idea so that EBV_{EXP} values adjusted for their reliability in the IMP country (EBV^*_{EXP}) are used, i.e.

$$EBV^*_{EXP} = (EBV_{EXP} - \overline{EBV}_{EXP}) * REL_{IMP} \quad [3]$$

where $\overline{EBV}_{EXP}$ is the average of EBVs in the EXP country and REL_{IMP} is the reliability of EBV in the IMP country. Then, using EBV^*_{EXP} in Equation (1) a modified slope (b^*) will be calculated from:

$$EBV_{IMP} = a + b^* \, EBV^*_{EXP} + \varepsilon \qquad [4]$$

In Wilmink *et al.* (1986) a modified intercept (a^*) is also calculated as:

$$a^* = \overline{EBV}_{IMP} - b^* \, \overline{EBV}_{EXP} \qquad [5]$$

Having calculated these parameters, the new adjusted converted breeding value (CBV^*_{IMP}) is predicted as:

$$CBV^*_{IMP} = a^* + b^* \, EBV_{EXP} \qquad [6]$$

For single trait regression analyses in the present study every Swedish trait was regressed on all USA traits one at a time.

Multiple trait regression analysis – Extension of single trait to multiple trait regression is straightforward and involves only a change of b or b^* from the above equations to be expressed as vectors instead of scalar.

Simple-MACE analysis – VanRaden (2001) suggested that given the EBVs and RELs for an animal, its sire and its dam (and some genetic parameters of the trait(s) of interest, e.g. heritability) it is possible to estimate the values for any of these EBVs and RELs, should any of these animals receive new information, or any of them have missing values, and even the information on a correlated trait can be used. In the present study VanRaden (2001) results for animal model, extended to sire-maternal grandsire model (Fikse, Unpublished results), were used in the following manner.

a) Obtain a nationally updated breeding value in the importing country ($NUBV_{IMP}$). For the dependent variable three alternatives were considered ($EBV_{IMP} = 0$, $EBV_{IMP} = CBV^*_{IMP}$ or $EBV_{IMP} = EBV_{IMP}$);

b) Obtain a national updated breeding value in the EXP country ($NUBV_{EXP(i)}$), where i = 1 USA trait for single trait analyses, or i = 1, 2, or 3 USA traits for the multiple trait analyses;

c) Using $NUBV_{IMP}$ from Sweden and $NUBV_{EXP(i)}$ for 1 trait or 3 traits (for single and multiple trait analyses, respectively) from the USA obtain an internationally updated breeding value ($IUBV_{IMP}$) for the Swedish trait of interest;

d) Simple-MACE (converted) breeding values (SBV_{IMP}) were assigned as $SBV_{IMP} = IUBV_{IMP}$.

Choice of traits for multiple trait analyses – In order to avoid computational overload any single Swedish trait was regressed on a set of three USA traits. The choice of the USA traits was based on the correlations between these

traits and three different alternatives were used.

The first alternative was to choose the three USA traits that had the highest product moment (Pearson) correlations with the Swedish trait of interest. In the second alternative product moment correlations adjusted for reliabilities of EBVs, the so called Calo-type correlations (Calo *et al.,* 1973; Blanchard *et al.,* 1984) were used. Calo-type correlations were calculated as follows:

$$r_{Calo} = r_{EBV_{IMP}, EBV_{EXP}} * \frac{\sqrt{\sum REL_{IMP} * \sum REL_{EXP}}}{\sum (REL_{IMP} * REL_{EXP})} \quad [7]$$

In the third alternative the first USA trait was the trait with the highest Calo-type correlation with the Swedish trait. The second USA trait was the USA trait with the highest partial correlation coefficient (for Calo-type correlation) with the Swedish trait given the USA trait already chosen. The third USA trait was the trait with the highest partial correlation coefficient (for Calo-type correlation) with the Swedish trait given the two USA traits already chosen.

Validation

For each analysis the bulls with EBV_{IMP} and $EBV_{EXP(i)}$ were randomly divided into two parts. The first part was used to estimate a, a^*, b, b^*, $r_{EBV_{IMP}, EBV_{EXP}}$ and r_{Calo}. Then, these parameters were used on the second part to calculate CBV_{IMP}, CBV^*_{IMP}, SBV_{IMP}. These different converted values were then compared with the reported EBV_{IMP} through:

i) The correlation between estimated and predicted breeding values, $r_{EST-PRE}$; and

ii) The mean squared error (MSE) of prediction.

This process of dividing data in two parts was repeated 100 times. Summarizing all the different methods and alternatives used there are six methods to compare:

STRE: Single trait regression equation – Based on Equations 1-3;

STCE: Single trait conversion equation – Based on Equations 4-6;

STSM: Single trait Simple-MACE – Using only one USA trait in the Simple-MACE method to calculate SBV for one Swedish trait. Depending on the input variable for computation of $NUBV_{IMP}$ there are three variants of this method, i.e. STSM-0, STSM-C, STSM-E (corresponding to $EBV_{IMP} = 0$, $EBV_{IMP} = CBV^*_{IMP}$ or $EBV_{IMP} = EBV_{IMP}$);

MTRE, MTCE and MTSM: Multiple trait versions of the above three methods.

Results and Discussion

In order to facilitate interpretations the results for another Swedish trait, i.e. Somatic Cell Count (SCC) will also be presented. Choice of SCC is motivated by the facts that:

a) It is a trait of relatively low reliability among the Swedish traits;

b) It has a direct equivalent trait in the USA, namely SCS, with a high correlation; and

c) The USA trait of interest, SCS, has one of the lowest mean reliability values among USA traits.

The Swedish TMI, being a composite trait itself, had generally higher correlations with the USA composite traits such as Cheese Merit \$ and Net Merit \$ ($r_{Calo} = 0.78$ and 0.77, respectively, Table 2). The $r_{EST-PRE}$ values for STRE and STCE (Table 3) indicate that the difference between the two methods STRE and STCE was so little that almost all results for TMI were identical for up to two decimal points. The reason for the lack of difference between STRE and STCE can be attributed to the near perfect reliability (0.91 ± 0.002) that the specific group of bulls with data on the Swedish TMI had. This was not always the case for all Swedish traits and in some cases the adjustment for reliability (e.g. in the STCE method) had a positive effect leading to higher $r_{EST-PRE}$ values for STCE than for STRE. The size of $r_{EST-PRE}$ was generally, but not always, very close to the product moment correlation

that the Swedish TMI had with the USA trait. This can be seen for example between the Swedish TMI and Cheese Merit $ and Net Merit $ or between the Swedish SCC and the USA SCS.

For the Swedish TMI the STSM-0 method had higher prediction power than the simpler methods STRE and STCE in combination with some traits. For the Swedish trait SCC the STSM-0 method generally showed higher $r_{EST-PRE}$ values than either of STRE and STCE. It seems that whenever the correlation is high, simple regression analysis methods (STRE and STCE) do a fine job of predicting a future observation, while in cases where the correlation is low Simple-MACE methodology "maybe" able to compensate for the low correlation.

The alternative STSM-C, which used CBV^*_{IMP} instead of EBV_{IMP} showed the same general pattern, that is having higher $r_{EST-PRE}$ in combination with some traits and lower $r_{EST-PRE}$ in combination with other traits, in comparison with STRE, STCE and STSM-0.

Looking at the results in Table 3 a legitimate question is why are not Simple-MACE methods always better than simple regression methods? The key to this question lies in the unlocking the role of the reliabilities of EBVs reported for a bull's relatives and the correlation with the foreign trait. A good example to illustrate this is the result for the Swedish trait SCC and the USA trait SCS and the STSM-E method. In STSM-E the breeding value of the bull in Sweden, ENP_{IMP}, was used in evaluations. Naturally, if no other national and international information existed the updated breeding values, $NUBV_{IMP}$ and $INUBV_{IMP}$, would be still equal to ENP_{IMP} which lead the value of $r_{EST-PRE}$ to be equal to unity.

For choice of the USA traits used in multiple trait methods (Table 4) while the USA traits Cheese Merit $ and Net Merit $ were predominately used as the first and the second traits the choice of the third trait was very much influenced by the sampling effect when partial correlation coefficient was used.

Results of multiple trait analyses (Table 5) indicated that MTRE and MTCE were always,

though to varying degrees, better than STRE and STCE. For traits with low correlations the improvement in results achieved by MTRE and MTCE was very noticeable (results not shown). The three Simple-MACE methods had in most cases lower $r_{EST-PRE}$.

Conclusion

We firmly believe multiple trait methods within the Simple-MACE methodology should be preferred to the single trait and simple regression methodology. We maintain this opinion mainly because we consider all of the reported EBV_{IMP} values used in the present study subject to sampling variance which is prone to "natural" fluctuation. Should those bulls be evaluated again together with a new group of daughters or a new group of bulls, new EBV_{IMP} would be obtained for these bulls.

Acknowledgment

References

Goddard, M. 1985. A, method of comparing sires evaluated in different countries. *Livest. Prod. Sci. 13,* 321-331.

International Dairy Federation, 1981. IDF recommended procedure for international comparison of genetic merit of dairy cattle. *International Dairy Federation, A-Doc 64.*

Jorjani, H. 2003. An overview of validation issues in national genetic evaluation systems (N-GES). *Interbull Bulletin 30,* 49-55.

Powell, R.L. 1988. Comparison of procedures to convert sire evaluations between countries. *J. Dairy Sci. 71,* 1609-1615.

VanRaden, P.M. 2001. Methods to combine estimated breeding values obtained from different sources. *J. Dairy Sci. 84 (E. Suppl.),* E47-E55.

Wilmink, J.B.M., Meijering, A. & Engel, B. 1986. Conversion of breeding values for milk from foreign populations. *Livest. Prod. Sci. 14,* 223-229.

Table 2. Mean reliability of the USA trait and their correlation with the Swedish traits total merit index (TMI) and somatic cell count (SCC). REL= reliability, SE=standard error, N=number of bulls with reported breeding values for the trait combination, PMC=product moment correlation between reported breeding values, Calo=correlation between reported breeding values adjusted for reliability. Values are averages of 100 replicates.

| | | | | Swedish Trait | | | | | | |
| | REL | | | TMI | | | SCC | | | |
USA Trait	Mean	SE	N	PMC	SE	Calo	N	PMC	SE	Calo
Milk Yield	0.93	0.004	54.4	0.30	0.13	0.32	84.8	-0.01	0.11	-0.01
Fat Yield	0.93	0.004	54.4	0.48	0.12	0.50	84.8	0.13	0.11	0.15
Protein Yield	0.93	0.004	54.4	0.55	0.12	0.58	84.8	-0.03	0.11	-0.04
Somatic Cell Score	0.73	0.015	54.4	-0.30	0.13	-0.32	84.8	-0.79	0.07	-0.91
Angularity	0.91	0.004	53.9	0.02	0.14	0.02	83.2	-0.07	0.11	-0.09
Udder Depth	0.91	0.004	53.9	0.31	0.13	0.33	83.2	0.36	0.10	0.41
Rear Udder view	0.91	0.004	53.9	0.18	0.14	0.19	83.2	0.24	0.11	0.28
Chest Width	0.91	0.004	53.9	0.15	0.14	0.16	83.2	0.26	0.11	0.30
Body Depth	0.91	0.004	53.9	0.13	0.14	0.14	83.2	0.20	0.11	0.23
OCS	0.92	0.004	53.9	0.36	0.13	0.38	83.2	0.32	0.10	0.37
OFL	0.91	0.004	53.9	0.33	0.13	0.35	83.2	0.14	0.11	0.17
OUS	0.91	0.004	53.9	0.38	0.13	0.40	83.2	0.34	0.10	0.39
Productive Life	0.70	0.014	54.4	0.49	0.12	0.52	84.8	0.32	0.10	0.38
Net Merit $	0.85	0.009	58.4	0.72	0.09	0.77	88.4	0.32	0.10	0.37
Fluid Merit $	0.85	0.009	58.4	0.54	0.11	0.57	88.4	0.31	0.10	0.36
Cheese Merit $	0.85	0.009	58.4	0.73	0.09	0.78	88.4	0.30	0.10	0.35

Table 3. Results of single trait analyses. Correlation of estimated and predicted (converted) breeding values for the two Swedish traits TMI and SCC based on 29 USA traits. STRE and STCE=Single trait conversion equations without and with adjustment for reliability, respectively, STSM-0, STSM-C and STSM-E=single trait simple MACE where the Swedish breeding values were ignored, were set to the converted values or were used as they are, respectively. Values are averages of 100 replicates.

| | | | Swedish Trait | | | | | | | |
| | | | TMI | | | | | SCC | | |
USA Trait	STRE	STCE	STSM-0	STSM-C	STSM-E	STRE	STCE	STSM-0	STSM-C	STSM-E
Milk Yield	0.29	0.30	0.25	0.27	1.00	-0.06	-0.06	0.23	0.13	1.00
Fat Yield	0.47	0.47	0.25	0.38	1.00	0.09	0.10	0.26	0.22	1.00
Protein Yield	0.56	0.56	0.41	0.49	1.00	-0.04	-0.04	0.22	0.14	1.00
Somatic Cell Score	0.27	0.27	0.08	0.19	1.00	0.79	0.79	0.06	0.72	0.94
Angularity	-0.08	-0.08	0.26	0.04	1.00	0.00	-0.02	0.21	0.02	1.00
Udder Depth	0.30	0.30	0.31	0.18	1.00	0.35	0.35	0.25	0.16	1.00
Rear Udder view	0.13	0.14	0.24	0.07	1.00	0.24	0.24	0.26	0.07	1.00
Chest Width	0.07	0.07	0.25	0.09	1.00	0.25	0.25	0.26	0.08	1.00
Body Depth	0.09	0.09	0.26	0.09	1.00	0.20	0.20	0.25	0.06	1.00
OCS	0.35	0.35	0.34	0.23	1.00	0.32	0.32	0.24	0.11	1.00
OFL	0.32	0.32	0.28	0.25	1.00	0.09	0.10	0.21	0.02	1.00
OUS	0.37	0.37	0.35	0.23	1.00	0.34	0.34	0.22	0.11	1.00
Productive Life	0.50	0.50	0.12	0.40	1.00	0.32	0.32	0.20	0.29	1.00
Net Merit $	0.72	0.72	0.39	0.73	1.00	0.31	0.31	0.26	0.27	1.00
Fluid Merit $	0.53	0.53	0.29	0.46	1.00	0.29	0.29	0.25	0.27	1.00
Cheese Merit $	0.73	0.73	0.41	0.74	1.00	0.29	0.29	0.26	0.26	1.00

Table 4. Number of times each USA trait has been used together with the Swedish TMI in different multiple trait analyses in 100 replicates.

	Choice of traits								
	Product moment correlation			Calo-type correlation			Partial correlation coefficient		
	1st	2nd	3rd	1st	2nd	3rd	1st	2nd	3rd
Milk Yield	0	0	0	0	0	0	0	0	30
Fat Yield	0	0	8	0	0	5	0	0	15
Protein Yield	1	0	42	0	0	35	0	0	22
SCS	0	0	0	0	0	0	0	0	1
Angularity	0	0	0	0	0	0	0	0	8
Udder Depth	0	0	0	0	0	0	0	0	1
Rear Udder view	0	0	0	0	0	0	0	0	3
Chest Width	0	0	0	0	0	0	0	0	2
Body Depth	0	0	0	0	0	0	0	0	1
OCS	0	0	1	0	0	1	0	0	0
OFL	0	0	2	0	0	2	0	0	1
OUS	0	0	1	0	0	1	0	0	0
Productive Life	0	0	18	0	0	19	0	0	3
Net Merit $	19	80	1	19	81	0	19	74	0
Fluid Merit $	0	0	27	0	0	37	0	7	13
Cheese Merit $	80	20	0	81	19	0	81	19	0

Table 5. Results of multiple trait analyses for the two Swedish traits TMI and SCC. MTRE and MTCE=Multiple trait conversion equations without and with adjustment for reliability, respectively, MTSM-0, MTSM-C and MTSM-E=Multiple trait simple MACE where the Swedish breeding values were ignored, were set to the converted values or were used as they are, respectively. Values are averages of 100 replicates.

Choice Of USA Trait	Swedish Trait									
	TMI					SCC				
	MTRE	MTCE	MTSM-0	MTSM-C	MTSM-E	MTRE	MTCE	MTSM-0	MTSM-C	MTSM-E
$r_{EST-PRE}$										
PMC	0.74	0.74	0.67	0.40	0.92	0.79	0.79	0.21	0.19	0.30
Calo	0.73	0.73	0.66	0.39	0.91	0.78	0.79	0.21	0.20	0.31
Partial	0.56	0.57	0.59	0.30	0.84	0.76	0.77	0.19	0.16	0.32
SE of $r_{EST-PRE}$										
PMC	0.09	0.09	0.10	0.12	0.05	0.07	0.07	0.10	0.11	0.10
Calo	0.09	0.09	0.10	0.12	0.05	0.07	0.07	0.10	0.11	0.10
Partial	0.11	0.11	0.11	0.12	0.06	0.07	0.07	0.10	0.11	0.10