Mace
The Relative Importance of Information Sources

***Bert Klei*[1], *Kent Weigel*[2], *and Tom Lawlor*[1]**
[1]Holstein Association USA, Inc., Brattleboro, Vermont, USA
[2]University of Wisconsin, Madison, Wisconsin, USA

Introduction

One of the concerns about Mace is that it puts too much emphasis on the parent average (**PA**) in the importing country. As a result of this, the argument goes, sons of sires that are over-evaluated in an importing country will always receive inflated evaluations. Another concern is that no matter how many daughters a bull has in an exporting country the weight on his importing country's PA stays the same. This paper will show a method to investigate the relative weights given to various sources of information that impacts a bull's Mace.

To assess the impact of all sources of information on a bull's evaluation one would have to take a look at the row of the inverse of the mixed model equations associated with bull j in country i. Since this is not a trivial task and there are many information sources that have a small influence, this paper will focus on only those equations of the mixed model pertaining to an individual bull. Methods from this paper can easily be extended to deal with more complex situations.

In this paper the focus will be on a two country situation where a bull has PA information as well as deregressed proofs (**DP**) based on a number of daughters in one or both of the countries.

Method

The following symbols and matrices are defined as follows.

h_i^2 : heritability in country i.

n_{ij} : number of daughters of bull j in country i.

DP_{ij} : deregressed proof for bull j in country j.

u_{ij} : breeding value of bull j in country i.

$g_{mgd,i}$: maternal granddam phantom parent group effect in country i.

c_i : effect of country i.

PA_{ij} : the parent average of bull j in country i $\left(PA_{ij} = .5 \times u_{sire,j} + .25 \times u_{mgs,j} + .25 \times g_{mgd,j}\right)$

d_j : constant resulting from relationship information for bull j ($d_j = 11/16, 12/16, 15/16, 16/16$ for known sire and mgs, unknown mgs, unknown sire, and both sire and mgs unknown, respectively).

r_i : residual variance in country i

G_o : genetic (co-)variance structure among the countries with elements $g_{ii'}$

$\rho_{ii'}$: genetic correlation between country i and i'.

α_i : $\dfrac{g_{ii}}{r_i} = \dfrac{h_i^2}{4 - h_i^2}$

If one extracts from the mixed model equations the three rows pertaining to bull j, these can be written as:

$$\left[\begin{pmatrix} n_{1j}r_1^{-1} & 0 \\ & n_{2j}r_2^{-1} \end{pmatrix} + \begin{pmatrix} d_j^{-1}g^{11} & d_j^{-1}g^{12} \\ & d_j^{-1}g^{22} \end{pmatrix}\right]\begin{pmatrix} u_{1j} \\ u_{2j} \end{pmatrix} - \begin{pmatrix} d_j^{-1}g^{11} & d_j^{-1}g^{12} \\ & d_j^{-1}g^{22} \end{pmatrix}\begin{pmatrix} PA_{1j} \\ PA_{2j} \end{pmatrix}$$
$$= \begin{pmatrix} n_{1j}r_1^{-1} & 0 \\ & n_{2j}r_2^{-1} \end{pmatrix}\begin{pmatrix} DP_{1j} - c_1 \\ DP_{2j} - c_2 \end{pmatrix}$$

Which can be expressed as:

$$\begin{pmatrix} 1 + d_j n_{1j}\alpha_1 & \rho^{12} \\ & 1 + d_j n_{2j}\alpha_2 \end{pmatrix}\begin{pmatrix} u_{1j} \\ u_{2j} \end{pmatrix} = \begin{pmatrix} d_j n_{1j}\alpha_1 & 0 \\ & d_j n_{2j}\alpha_2 \end{pmatrix}\begin{pmatrix} DP_{1j} - c_1 \\ DP_{2j} - c_2 \end{pmatrix} + \begin{pmatrix} 1 & \rho^{12} \\ & 1 \end{pmatrix}\begin{pmatrix} PA_{1j} \\ PA_{2j} \end{pmatrix}$$

If one writes these equations as $(A+B)\left(u_{.j}\right) = \left(A \;\vdots\; B\right)\begin{pmatrix} DP_{.j} - c_{.} \\ \cdots \\ PA_{.j} \end{pmatrix}$. Weights on each of the three DPs and PAs can be obtained from:

$$\left(u_{.j}\right) = (A+B)^{-1}\left(A \;\vdots\; B\right)\begin{pmatrix} DP_{.j} - c_{.} \\ \cdots \\ PA_{.j} \end{pmatrix} = (A+B)^{-1}A\left(DP_{.j} - c_{.}\right) + (A+B)^{-1}B\left(PA_{.j}\right).$$

Which in scalar notation results in $u_{ij} = \sum_i b_{iD}\left(DP_{ij} - c_i\right) + \sum_i b_{iP}PA_{ij}$ where b_{iD} and b_{iP} are the regression coefficients in country i for the DPs and PAs , respectively.

The relative importance of each components b_{ik}^{rel} (k is D or P) can then be determined by calculating $b_{ik}^{rel} = \left|b_{ik}\right| \times \frac{\sum_{i'=1}^{2} b_{i'k}}{\sum_{i'=1}^{2}\left|b_{i'k}\right|}$ for each of the regression coefficients.

Results and Discussion

To simplify the discussion, results will only be presented from one country's perspective. The regression coefficients are expressed so that they always sum up to 1 (Table 1 through Table 3). Observe from these Tables that the regression coefficient for DP_2 is always of the same size but of opposite magnitude as the one for PA_2.

Let's first examine the example of 0 daughters in country 1 and 100 daughters in country 2 (Table 1) to illustrate the calculation of the relative importance of DP and PA. In this case $\sum_{i'} b_{i'DP} = \sum_{i'}\left|b_{i'DP}\right| = .74$, while $\sum_{i'} b_{i'PA} = .26$ and $\sum_{i'}\left|b_{i'PA}\right| = 1.74$. From these results it follows that $b_{1DP}^{rel} = 0$, $b_{2DP}^{rel} = .74$, $b_{1PA}^{rel} = .15$, and $b_{2PA}^{rel} = .11$. Looking at the actual regression coefficients the impression is created that the importance of PA in the importing country is much larger than what is actually the case.

Table 1. Regression coefficients and relative importance of information sources in country 1 for a two country situation for different daighters distibution and fixed heritabilities and correlation[1]

Number of Daughters		Regression Coefficients: Deregressed Proofs		Regression Coefficients: Parent Averages		Relative Importance (%): Deregressed Proofs		Relative Importance (%): Parent Averages	
1	2	1	2	1	2	1	2	1	2
0	0	.00	.00	1.00	.00			100.0	
100	0	.82	.00	.18	.00	82.1		17.9	
0	100	.00	.74	1.00	-.74		73.9	15.0	11.1
0	1000	.00	.88	1.00	-.88		88.1	6.3	5.6
50	50	.50	.31	.50	-.31	50.0	31.3	11.5	7.2

[1] fixed parameters: $h_1^2 = .25$, $h_2^2 = .25$, $correlation = .90$

A closer examination of Table 1 reveals the importance of the number of daughters. When no daughters are present in either country, all the information is coming from PA in the country 1. A very familiar situation is presented in the situation of 100 daughters in country 1 and 0 in country 2 where as a result 18% of the evaluation is based on PA and 82% on DP and no information is provided from those countries without any daughters. When there are daughters in another country, and no daughters in the importing country, three sources of information are important. Also from Table 1 it can be observed that an increase in number of daughters in the foreign country will put more emphasis on the DP in the foreign country with a reduction in emphasis on PA from both countries. Note that with daughters in a foreign country the emphasis on PA is divided between both countries. The importance of good pedigree information can be seen from this example, since the PA from both countries are contributing to a bull's Mace.

Table 2 shows the influence of different heritabilities. As expected an increase in heritability shifts the emphasis from PA to DP. Also from Table 2 it can be seen that higher heritabilities in the foreign country puts more emphasis on DP in that country.

Table 2. Regression coefficients and relative importance of information sources in country 1 for a two country situation for different heritabilities and fixed number of daughters and correlation[1]

Heritabilities		Regression Coefficients: Deregressed Proofs		Regression Coefficients: Parent Averages		Relative Importance (%): Deregressed Proofs		Relative Importance (%): Parent Averages	
1	2	1	2	1	2	1	2	1	2
.25	.25	.50	.31	.50	-.31	50.0	31.3	11.5	7.2
.25	.40	.45	.39	.55	-.39	45.1	39.2	9.2	6.6
.40	.25	.62	.24	.38	-.24	62.5	23.5	8.6	5.4
.40	.40	.58	.30	.42	-.30	57.8	30.1	7.1	5.0

[1] fixed parameters: $n_1 = 50$, $n_2 = 50$, $correlation = .90$

Table 3. Regression coefficients and relative importance of information sources in country 1 for a two country situation for different heritabilities and fixed number of daughters and heritabilities[1]

	Regression Coefficients				Relative Importance (%)			
Correlation	Deregressed Proofs		Parent Averages		Deregressed Proofs		Parent Averages	
	1	2	1	2	1	2	1	2
.90	.50	.31	.50	-.31	50.0	31.3	11.5	7.2
.80	.56	.25	.44	-.25	56.0	24.5	12.5	7.0
.70	.60	.19	.40	-.19	60.2	19.4	13.7	6.7
.60	.63	.15	.37	-.15	63.2	15.4	15.1	6.3

[1] fixed parameters: $n_1 = 50$, $n_2 = 50$, $h_1^2 = .25$, $h_2^2 = .25$

With a lower correlation the number of effective daughters over the two populations decreases, which results in a shift in emphasis from DP to PA (Table 3). Also when the correlation reduces, the relative emphasis of the information from the foreign country reduces and subsequently one observes a shift in relative importance from the exporting country to country importing country.

In Table 4 a three country situation is given. This is a simple extension of the method described. In this table a results are given for a situation in which 300 daughters are divided over 3 countries with fixed heritabilities and correlations. As can be seen from Table 4 the best way to maximize the Mendelian Sampling contribution is to have daughters located in different countries. This will be counter intuitive to some people believing that if one is only interested in country 1, then the most efficient progeny test would be to put all daughters in country 1. Distributing the daughters also guarantees that an evaluation does not become solely reliant on PA in country 1 only. For the parameters used in this example, a daughter distribution of 230, 25, and 45 over the three countries respectively is optimal. If daughters could only be in the two foreign countries, then 215 in country 2 and 85 in country 3 would be optimal. Results in this Table show some of the interactions that occur between correlations and heritabilities when determining relative importance.

Table 4. Regression coefficients and relative importance of information sources in country 1 for a three country situation for different distributions of 300 daughters and fixed heritabilites and correlation[1]

			Relative Importance (%)							
Number of daughters			Deregressed Proofs				Parent Averages			
1	2	3	1	2	3	Total	1	2	3	Total
300	0	0	94.4			94.4	5.6			5.6
200	100	0	78.9	15.6		94.5	3.2	2.3		5.5
200	0	100	83.2		11.7	94.9	3.0		2.1	5.1
100	200	0	60.1	32.4		92.5	4.2	3.4		7.5
100	100	100	61.1	20.4	12.0	93.5	3.6	1.9	1.1	6.5
100	0	200	69.3		22.8	92.1	4.5		3.4	7.9
0	300	0		83.9		83.9	8.8	7.3		16.1
0	200	100		64.3	22.0	86.3	7.4	4.7	1.6	13.7
0	100	200		49.0	35.6	84.6	8.3	4.1	3.0	15.4
0	0	300			76.1	76.1	13.6		10.3	23.9
230	25	45	83.9	4.7	6.6	95.3	2.8	.8	1.1	4.7
0	215	85		66.1	20.2	86.3	7.3	4.9	1.5	13.7

[1] fixed parameters: $h_1^2 = .30$, $h_2^2 = .25, h_3^2 = .35$, $correlation = \begin{pmatrix} 1 & .90 & .80 \\ & 1 & .85 \\ & & 1 \end{pmatrix}$

Conclusions

The procedure shown in this paper shows a method in which to assess the relative importance of information sources on Mace. As opposed to looking at the regression coefficients, results from the method described are easy to interpret and agree with expectations.

Acknowledgments

The authors would like to thank the National Association of Animal Breeders for financial support for this project.