Interbull Checks of International Breeding Values

Eva Hjerpe & Jette H. Jakobsen
Interbull Centre, Box 7023, SE-750 07 Uppsala, Sweden

Introduction

Interbull routinely verifies the national evaluations results before it is included in the international genetic evaluation. The verification tools used are described in by Jakobsen & Hjerpe (2006) and Klei *et al.* (2002).

National evaluation units are motivated to release national breeding values of high quality, and similarly Interbull is motivated to release international breeding values of high quality. As a MACE evaluation involves many different parameters, it requires sophisticated procedures to detect potential problems.

The aim of this paper is to describe three methods used by Interbull Centre, to verify international breeding values.

Comparison of current and previous international breeding values

The first method used to verify the international breeding values is the comparison of current and previous international evaluations. This comparison is done for each country scale involving all bulls with an international breeding value. The intenational breeding values are calculated using the MACE model according to Schaeffer (1994) and reliabilities of international breeding values are approximated according to Harris and Johnson (1998).

To check the change between the current and previous international evaluation, a statistic (STAT) is calculated for each bull and country scale. The STAT value relates the change in international breeding values to the reliability of current and previous evaluation and is calculated as:

$$STAT_{bi} = \frac{|\, u_{bit} - u_{bi(t\text{-}1)} \,|}{\sqrt{\mathrm{var}\left(u_{bit} - u_{bi(t\text{-}1)}\right)}} \quad ,$$

where: u_{bit} and $u_{bi(t-1)}$ are the evaluations for bull b on country scale i on time t and t-1, resp.

$$\mathrm{var}\left(u_{bit} - u_{bi(t\text{-}1)}\right) = REL_{bit} \cdot \sigma^2_{u(t-1)} - 0.80 \cdot REL_{bi(t-1)} \cdot \sigma^2_{u(t-1)}$$

where REL_{bit} and $REL_{bi(t-1)}$ are the reliability of the international breeding values for bull b on country scale i on time t and t-1, resp., and $\sigma^2_{u(t-1)}$ is the (REML) estimated sire variance on time t-1. $\mathrm{var}(u_{bit} - u_{bi(t-1)})$ is computed according the same principles applied in the verification of national breeding (Klei *et al.,* 2002).

Bulls with a STAT value higher than 5 are printed and further investigated as recommended by Klei *et al* 2002. In the output (see Table 1 for an example) there is information about bull ID and country scale in addition to the STAT value and its components.

At this stage in the international genetic evaluation process, there is usually enough information about changes in the national evaluations to make a good judgment of the high values. If there is a questionable change, for instance one country with many high STAT values but with no changes in the national evaluation, the country in question is contacted and asked to investigate the reason for the change. Many bulls with a high STAT value is usually a result of changes in the model or change of genetic correlations.

Table 1. Comparison of current and previous international breeding value.

STAT	COU	BULLID	REL_CUR	REL_PRV	PRF_CUR	PRF_PRV
5.5	Cou 1	111	3	3	-18.12	-26.99
5.6	Cou 1	222	13	15	-12.15	-24.46
5.9	Cou 1	333	17	17	-9.31	-23.49
7.8	Cou 1	444	16	16	-10.00	-28.35

STAT=statistics, COU=country scale, BULLID=bull identification number, REL_CUR=current reliability, REL_PRV=previous reliability, PRF_CUR=current international breeding value, PRF_PRV=previous international breeding value.

To summarize the comparison between current and previous international breeding values, Pearson correlations, regressions, numbers of bulls with extreme changes and numbers of bulls with > 10% decrease in reliability are calculated (Table 2). These checks are done for each country scale considering all bulls with an international breeding value.

Both correlation and regression are preferred to be close to one but not lower than 0.97. With no changes in the national or international evaluation, the correlation and regression are expected to be one.

In Table 2 we observe correlation and regression of 0,988 and 0,838, respectively, for GBR. This low correlation and regression is due to the difference in genetic evaluation models for the UK in both evaluations.

In Table 2 we can also observe a regression of 1.034 for DEU. Often a regression that deviates substantially from one can be related to a change in the REML estimate of sire standard deviation. To pin point the specific reason for a questionable change is not straightforward as many parameters are taken into account in the MACE model.

Number of extreme changes are number of bulls with a STAT value equal or larger than 5. Changes in national and international evaluation as well as the size of population are taken into consideration when evaluating the number of bulls. During a test run, when changes in the national /international evaluations are allowed, changes are allowed to be higher than during a routine run.

The number of extreme changes also differs within trait group. In the example in Table 2, the number of bulls with extreme changes for milk on the DEU scale is 543. In this case, the number of bulls with extreme changes on the DEU scale were 63 for fat and 682 for protein (not shown).

The total change was a result of other countries changing their national evaluation and the difference within trait group is due to different heritabilities. Another aspect affecting the comparison is the number of imported bulls. International breeding values for a country with many imported bulls are often more sensitive to changes compared to countries with few imported bulls.

Table 2.

```
SUMMARY OF COMPARISONS PER COUNTRY SCALE
 ========================================

CURRENT VS PREVIOUS INTERNATIONAL EVALUATION
  Pearson correlation:
     CAN   FIN   NOR   SWE   USA   NZL   AUS   GBR   DNK   DEU   ZAF   EST
   0.996 0.996 0.996 0.996 0.996 0.996 0.995 0.988 0.996 0.996 0.996 0.996

  Regression cur|prv:
     CAN   FIN   NOR   SWE   USA   NZL   AUS   GBR   DNK   DEU   ZAF   EST
   0.981 0.998 0.991 0.984 0.996 0.991 1.005 0.838 0.995 1.034 0.997 0.965

  Number of extreme changes:
     CAN   FIN   NOR   SWE   USA   NZL   AUS   GBR   DNK   DEU   ZAF   EST
       1     2     2     2     2     2     1     3     2   543     1     2

  Number of bulls with >10% decrease in reliability:
     CAN   FIN   NOR   SWE   USA   NZL   AUS   GBR   DNK   DEU   ZAF   EST
      69   100   105    75   101    89    62    98    62    97    83   101
```

Comparison of current International proofs across standardized country scales

The second method compares each bull's international breeding values across standardized country scales. This method was implemented to detect problems with wrong direction signs of genetic correlations and it is also a way to highlight bulls that might have been preferentially used in one country. With high correlations between countries a bulls breeding value is not expected to differ much from one country scale to another. If the breeding value is high in one country the breeding value is expected to be high in another country.

Breeding values are first standardized to a mean of 100 and a standard deviation of 10. For each bull the 'home country' is determined as the country scale with the highest reliability. The international breeding values for that bull on all other country scales are compared with the international breeding value on the 'home country' scale, and a STAT value is calculated:

$$STAT_{bi} = \frac{\left| u_{std_{bit}} - u_{std_{bdt}} \right|}{15.0 / \left(rG_{i,d} \cdot REL_{bit} \right)}$$

where u_{std_bit} and u_{std_bdt} are the standardized international breeding values for bull b on country scale i and d (home country) on time t, respectively, $rG_{i,d}$ the genetic correlation between country i and d, and REL_{bit} the reliability of the international breeding value of bull b on country scale i and time t.

If the STAT value exceeds one then the bull is printed (Table 3). A country for which many bulls with a high STAT value are printed is investigated according to the same procedure as for the previous checks.

The listing of individual bulls with large differences in international breeding values across country scales is followed by summary per country scale (Table 4). The number of large deviations is the number of bulls with a STAT value larger than 1.

Table 3. Comparison of current International proofs across standardized country scales.

STAT	COU	BULLID	REL_FOR	REL_DOM	PRF_FOR	PRF_DOM	COU_DOM
1.02	Cou1	111	99	99	87.33	107.90	Cou7
1.06	Cou2	222	85	99	86.55	111.50	Cou7
1.44	Cou3	333	99	99	82.49	111.50	Cou7
1.13	Cou4	444	99	99	93.68	116.38	Cou7
1.05	Cou5	555	95	99	111.97	92.57	Cou8
1.07	Cou6	666	98	99	115.52	96.30	Cou8

STAT=statistics, COU=country scale, BULLID=bull identification number, REL_FOR=reliability foreign country, REL_DOM=reliability domestic country, PRF_FOR=foreign breeding value, PFR_DOM=domestic breeding value, COU_DOM=domestic country.

Table 4.

```
SUMMARY OF COMPARISONS PER COUNTRY SCALE
 ========================================

STANDARDIZED INTERNATIONAL SCALES
  Number of large deviations:
  CAN   FIN   NOR   SWE   USA   NZL   AUS   GBR   DNK   DEU   ZAF   EST
   0     0     0     0     0     1     0     0     0     0     0     0
```

Comparison of International and National breeding values

The third method compares the current national with the international breeding value for each bull (Table 5). The method utilizes the difference between national and international reliabilities to establish the criteria for probable changes in breeding values. The national reliability is approximated using the EDCs provided by the countries.

The statistics (STAT) is calculated as:

$$STAT_{bi} = \frac{\left| n_{bit} - u_{bit} \right|}{\sqrt{\text{var}\left(n_{bit} - u_{bit} \right)}} \ ,$$

where n_{bit} and u_{bit} are the national and international breeding value of bull b in country i on time t, resp.

$$\text{var}\left(n_{bit} - u_{bit} \right) = NREL_{bit} \cdot \sigma^2_{ut} - 0.8 \cdot REL_{bit} \cdot \sigma^2_{ut} \ ,$$

where $NREL_{bit}$ and REL_{bit} are the reliability of the national and international breeding values for bull b on country scale i on time t, resp., and σ^2_{ut} is the (REML) estimated sire variance on time t.

$$NREL_{bit} = \frac{REL_{bit\,(dau)} + REL_{bit\,(PA)} - 2 \cdot REL_{bit\,(dau)} \cdot REL_{bit\,(PA)}}{1 - REL_{bit\,(dau)} \cdot REL_{bit\,(PA)}}$$

where $REL_{bit}(PA)$ reflects the reliability of parent average (fixed for all animals assuming reliability of 90% of both sire and maternal grandsire), and $REL_{bit}(dau)$ is the reliability of bull b based on daughters for country i on time t, computed as:

$$REL_{bit\,(dau)} = \frac{EDC_{bit}}{EDC_{bit} + \left(4 - h^2_{it} \right) \big/ h^2_{it}} \ ,$$

where EDC_{bit} is the number of EDC of bull b for country i on time t, and h^2_{it} is the heritability for country i on time t.

A STAT value larger than 2 is printed (Table 5) and investigated.

To get an overview of the agreement between national and international breeding values for a country correlation, regression and number of large deviation is calculated (Table 6). The number of large deviations is the number of bulls with a STAT value higher than 2. A more strict STAT value is used as in the comparison between subsequent international evaluations reflecting that large differences between national and international breeding values are suspicious. The rules are otherwise the same as for comparison between the international current and previous evaluation. In the interpretation of the results of this comparison the population size and structure for each country are considered. The difference between national and international breeding values is usually larger for countries with many relatively many common bulls (e.g. due to imports, joint progeny testing).

Table 5. Comparison of current international and national proofs.

STAT	COU	BULL ID	REL_INT	REL_NAT	PRF_INT	PRF_NAT
2.02	Cou1	111	77	49	8.040	1.950
2.10	Cou2	222	88	68	-56.700	-34.000
3.77	Cou1	333	84	63	9.170	-1.570
2.18	Cou3	444	86	58	-9.490	16.000
2.09	Cou4	555	68	56	14.410	7.500
2.00	Cou5	666	82	65	-4.300	6.120
2.09	Cou6	777	75	45	9.520	3.0950

STAT= statistics, COU=country, BULLID=bull identification number, REL_INT=international reliability, REL_NAT=national reliability, PRF_INT=international breeding value, PRF_NAT=national breeding value

Table 6.

```
SUMMARY OF COMPARISONS PER COUNTRY SCALE
 ========================================


NATIONAL VS INTERNATIONAL EVALUATION
  Pearson correlation:
CAN   FIN     NOR   SWE   USA   NZL   AUS   GBR   DNK   DEU   ZAF   EST
0.998 0.998 1.000  0.999 0.984 0.997 0.994 0.998 1.000 0.999 0.981 0.992

  Regression int|nat:
    CAN   FIN   NOR   SWE   USA   NZL   AUS   GBR   DNK   DEU   ZAF   EST
   0.998 0.998 1.001 0.998 1.000 0.998 1.003 0.996 0.999 0.997 0.991 1.002

  Number of large deviations:
     CAN   FIN   NOR   SWE   USA   NZL   AUS   GBR   DNK   DEU   ZAF   EST
       0     2     0     0     0     0     0     0     0     0     0     0
```

Conclusions

Interbull routinely verifies national genetic evaluations results prior to inclusion in the Interbull evaluations with the aim to identify potential problems. Similarly, this paper described three different methods for verification of international genetic evaluation results that Interbull routinely performs prior to release of international evaluations results. These checks verify that international breeding values are consistent across country scales, that international breeding values are consistent with national breeding values, and that current and previous international breeding values are consistent. Statistics are computed both on population and individual level. The observations from the verification of international breeding values must be put in relation to the changes in the national and international genetic evaluations. It is important verify outgoing data both on the national and international level to assure a constant and high quality of international evaluations.

References

Harris, B. & Johnson, D. 1998. Information source reliability, method applied to MACE. *Interbull Bulletin 17,* 31-36.

Klei, B., Mark, T., Fikse, F. & Lawlor, T. 2002. A method for verifying genetic evaluation results. *Interbull Bulletin 29,* 178-182.

Shaeffer, L.R. 1994. Multiple-country comparison of dairy sires. *J. Dairy Sci. 77,* 2671-2678.