Interbull Validation Test for Genomic Evaluations

Esa Mäntysaari[1], Zengting Liu[2] and Paul VanRaden[3]
[1]Genetic Research, MTT Agrifood Research Finland, [2]VIT, Verden, Germany
[3]Animal Improvement Programs, USDA Agricultural Research Service, Beltsville, MD, USA
e-mail: esa.mantysaari@mtt.fi

Abstract

Genomic evaluations (GEBV) can be validated by comparing GEBVs of bulls in the youngest age classes to the daughter yield deviations (DYD) that the bulls will receive later. The GEBV are calculated from truncated data where the last four years of phenotypic data are removed. The test consists of weighted linear regression of DYD on GEBV. The intercept, regression coefficient and the coefficient of determination R^2 are interpreted to represent average bias, prediction bias, and the validation accuracy, respectively.

Key words: estimation of breeding values, genomic evaluation, validation, international evaluation, MACE

Introduction

In 2008, a genomics task force was set up by the Interbull Steering Committee to establish methodology for the use of genomic data in national and international evaluations (Banos, *et al.*, 2009). One of the expectations for the task force was to define methods for utilization of national genomic evaluations in across country evaluations and also to propose methods that guarantee the quality and usefulness of the national genomic evaluations as input data. The first mandate lead to development of genomic across country multi-trait evaluation (GMACE, Sullivan and VanRaden, 2010). The purpose of this paper is to describe an Interbull validation test that could be used for direct estimated genomic values (DGV) or genomically enhanced breeding value estimates (GEBV).

The genomic evaluations that are to be used in GMACE should be validated using a GEBV test. This assures that the GEBVs will contribute useful information to evaluations of animals in other countries, and that the international GEBVs of one country can be safely used in others. Moreover, the international GEBV validation will act as a quality assurance of evaluations, and gives a standard that certifies evaluations to have expected level of accuracy. Such evaluations can then be converted to other country scales, or directly used as estimated breeding values (EBV) in international trade.

Goals

In the first place, the GEBV test should evaluate the unbiasedness of the genomic evaluations. The GEBVs should accurately predict the breeding values, which later will be estimated using daughter information. The GEBV test is based on assumption that evaluations which are useful in practical selection process, are also best input for GMACE evaluation procedure. That means the GEBVs should reflect the genetic trend in the population, and that variation in GEBVs is consistent with the accuracy of the evaluations. Moreover, the accuracy of the genomic evaluations should be on average higher than what is achieved using pedigree information only. The GEBV validation test follows the same principals as the Interbull validation test method 3 (Boichard et. al. 1995). The method 3 tests the consistency of genetic trend by comparing the estimated breeding values (EBVs) of the bulls based on first crop daughters to their final EBVs that include also the new second crop daughters. In GEBV test the first evaluations are GEBVs or DGVs before the bull gets own daughters, and these are compared to the future progeny performance, measured as bull's daughter yield deviation (DYD).

Test data sets

The participating national genetic evaluation center (NGEC) shall prepare two data sets of breeding values for the GEBV test. These data sets are referred to as *full data set* and *reduced data set*. Both data sets must be on the same genetic base. The estimation of trend is based on change in breeding value estimates, thus the values of animals with reliable daughter information should be about the same in both data sets. Countries may submit relative breeding values, but test results are easier to interpret if they are centered so that animals considered as in the country's current genetic base have zero mean in both the full and the reduced data. If the country submits relative breeding values, the standardization of EBVs in both data sets should be done using the variance standardization coefficient of the base in full data set.

Full data set

This data set will consist of the data that the NGEC is submitting to GMACE evaluations. For the GEBV testing the data should include:

1. Conventional genetic evaluation information: EBV, EDC, r^2_{EBV}. The EDC (number of effective daughter contributions) and reliability r^2_{EBV} are useful in detecting inconsistencies in genomic information. In the GEBV test the EDC is used to derive the weighting factor for DYD. The EBV are to be used to estimate the amount of pre-selection among the genotyped bulls.

2. Genomic evaluation information: GEBV, GEDC, r^2_{GEBV}. The GEDC (genomic effective daughter contribution) are used in possible deregression needed for GMACE. If the GEDC is not available, it can be approximated as GEDC = k * r^2_{GEBV} / (1- r^2_{GEBV}), where k= (4- h^2)/ h^2. The method of estimation of GEDC (and/or r^2_{GEBV}) has to be reported in the Interbull GE[1] form.

3. Phenotypic information: DYD or de-regressed proofs (DRP). Optimally the test would require DYDs, but if unavailable the countries should de-regress the EBVs with best available method to make the daughter performance measures independent from the pedigree.

4. DGV reference group indicator: identifies the animals used in the reference group of estimation of DGV prediction equations. This is needed in GMACE to check the amount of information shared across countries in derivation of DGV.

In above the genomic evaluations are considered to be GEBVs which implies that the NGEC have combined information in the EBVs and DGVs. In principal the countries can submit to Interbull either EBVs or DGVs, but consistently the same indices that they publish.

Reduced data set

This data set should include the same information as the full data but calculated from reduced phenotypic data after deleting *ny* years of observations. To be consistent with Interbull Testing Method 3, the value of *ny* should be equal to 4 years. Hereafter the variables in reduced data are called the same as in full data, but with a subscript r (e.g. $GEBV_r$ for genomic evaluations from the reduced data).

The genomic prediction equations used to compute DGV_r should also be based on the reduced data. Moreover, if the country submits GEBVs, the EBVs used in combining should also be from the reduced data. The aim is to mirror the situation that has been 4 years ago. If available, the actual EBVs that were used in 4 years ago can be used as reduced data set. Then the NGEC should report to Interbull the estimate of the expected correlation R between estimation methods in the full data and the reduced data (Boichard *et al.*, 1995).

Suggested exceptions

a) The accuracy of DGV prediction equations depends on the size of the reference population. If the size of the reference population with reduced data is significantly smaller than the actual reference population

[1] Interbull GE forms to report the national genetic evaluation methods are available at www.interbull.org

using full data, the realized validation accuracy is much smaller than what would be true today. In most of the evaluations 4 years data cut off corresponds about 20% of the total number of genotyped bulls. We suggest adjusting the *ny* so that the size of the reference population in reduced data set would be between 20-25% of the size of the reference population, but that the number of test bulls (having EDC_r=0 and EDC>20) would be more than 150.

b) When re-estimation of EBVs using the reduced data set is impossible, the EBVs used for estimation of prediction equations can be extracted from the full data EBVs. These should be the bulls that could not have had daughters in reduced data. These bulls should be reported in the reduced data with EDC_r=0, and $GEBV_r$ with appropriate $GEDC_r$.

GEBV Validation test

The bias in the evaluations will be tested using a regression model:

$$Y = b_0 + b_1*GEBV_r + e \qquad [1]$$

where Y is the DYD (if not available, then DRP) of the test bulls that have EDC>20 and EDC_r =0. Either the EDC_i or the accuracy of the DYD, estimated as $w_i=EDC_i/(EDC_i+k)$, will be used as weights for Y of the bull i.

The expectation for the model effects are b_0=0.0 (average bias) and b_1=1.0 (prediction bias). However, in cases where the test bulls are not representative sample of all the bulls in their year classes, the expected value of b_1 is reduced. The new H_0 hypothesis is $b_1=\beta$, where the expected value β is approximated using the estimate of preselection on EBV and the estimated effect of selection on variances and covariances (see Appendix). The statistical significance of b_0 can be tested using t-test, and the for the b_1 the H_0 is accepted if the $|b_1-\beta| < 2*SE(b_1)$.

The validation reliability of the model is estimated as $R^2/\overline{w}$, i.e.

$$R^2_{validation} = R^2(1+k/\overline{EDC}),$$

where $\overline{EDC}$ is average EDC of the test bulls. This value is compared to the average r^2_{GEBVr} of the test bulls reported in reduced data set.

The improvement obtained using GEBV instead of pedigree index (EBV_r) will be estimated by comparing the R^2 from model [1] with the R^2 from the model:

$$Y = b_0 + b_1*EBV_r + e \qquad [2]$$

The R^2 from model [1] should be notably higher than the R^2 from model [2].

Discussion

The three previously used Interbull methods of validation of national genetic evaluations are mostly focusing on unbiased estimation of genetic trend. Wrong genetic trend in one country is known to hamper the accuracy of international evaluations. This is likely to be true with GMACE also, although the wrong trend in GEBVs is caused by other reasons than wrong trend in EBVs. If the level of GEBVs of young bulls in one country is under- or overestimated, this will reflect to GMACE estimates of their sires also. This will thereafter lead to wrong across country comparisons.

Another serious bias on GMACE will result from the selection based on GEBVs (Patry and Ducroucq, 2009). This selection will lead to underestimation of genetic trend in national evaluations, and is likely to be seen in genomic evaluations as well. The selection can be accounted by including the GEBVs of bulls culled into national evaluations. After this, however, the evaluations might not be useable in estimation of DGV prediction equations.

The validation method presented will mimic the true accumulation of information across time. In practice the GEBVs are used to select young bulls or bull calfs. After 3-5 years the bulls will receive daughter information, and it is desirable that the GEBVs would highly correlate to final EBVs. If the reduced data GEBVs are based on DGV prediction equations derived from full data bull subset (exception b in validation method), the test will not reflect true accuracy. Such GEBV (or

DGV) are derived from more accurate EBVs than those derived from true reduced data set. On the other hand, use of low accuracy bulls in the closest reference population age class might lead to DGVs that will predict better early lifetime records than performance in latter lactations.

The expectation of b_1 in model [1] is 1.0 only if the test bulls are representative sample of all the bulls in their comparison group. If the bulls genotyped have been selected based on their EBVs, the expectation of b_1 as well as the expected R^2 of the model are reduced. The reduction in expectation can be approximated using the formulas in Appendix. However, more experiences on the test are needed until the final acceptance or rejection thresholds can be set. In the current Interbull genetic trend validation method III a special rule is applied to small populations with uncertainty in trend estimates. Similar guidelines for handling small genotyping populations might need to be developed, if the proposed test is found not well suitable for small genotyping populations

Acknowledgements

Discussion and suggestions from the Interbull genomics task force and Interbull Technical Committee were much appreciated.

References

Boichard, D., Bonaiti, B., Barbat, A. & Mattalia, S. 1995. Three methods to validate the estimation of genetic trend for dairy cattle. *JDS 78l*, 431-437.

Falconer, D. S. & Mackay, T.F.C.M 1996. *Quantitative Genetics*, Longman House, Essex, England.

Patry, C. & Ducroucq, V. 2009. Evidence of a Bias in Genetic Evaluation due to Genomic Selection. *Interbull Bulletin 40,* 167-171.

Sullivan, P.G. & VanRaden, P.M. 2009. Development of genomic GMACE. *Interbull Bulletin 40,* 157-161.

VanRaden, P.M. & Sullivan, P.G. 2010. International genomic evaluation methods for dairy cattle. *Gen. Sel. Evol. 42:7.*

APPENDIX: Expected value of b_1 in GEBV validation test.

Expected value of b_1 is 1.0 only if the genotyped test bulls are representative sample of the bulls in the corresponding age classes. The selection based on EBVs will reduce the value of the b_1 and equally the value of R^2 of the model [1].

The effect of selection can be approximated by examining the mean EBV of the genotyped test bulls and the mean EBV of all test bull candidates (i.e. bulls with EDC > 20 but EDC_r=0). Based on difference of these means and the SD of EBVs of the test bulls the proportion of selective genotyping can be estimated. Using table values from quantitative genetics text books, (e.g. Falconer and Mackay, 1996) the selection difference can be expressed as a proportion selected p.

When the proportion selected is known, the expected value of the b_1 and the effect of the selection on R^2 of the test model can be estimated by approximation of the effect of selection on the variance of the selected trait and on the covariance between independent and dependent variable regression effect.

This can be done using, for example, following commands of the system R:

```
#  INPUT:
#  p = the fraction of animals not genotyped;
#        selected based on EBV
#  r2 = R2 from the validation test
#
#  OUTPUT:
#  the R2BeforeSelection =
#     "a guess" what the true R2 could be
#     given your estimated R2
# the CorrelationAfterSelection =
#     correlation after 1-p selection given the
#     input R2 would be true before selection
# the expected.b1 =
#     the expected value of the b1 (regression of EBV
#     on GEBV) if given the true  R2

      a <-  qnorm(p);  l <- dnorm(a)/(1-p); k < -l*(l-
a)
      v1 <- (1-k)
      v2 <- (1-k*r2)
      CorrelationAfterSelection <- sqrt(r2*v1/v2)
      R2BeforeSelection <- r2*v2/v1
      expected.b1 <- v1/v2
```

Example of effect of selection on b1 and R^2

The difference between the mean of test bulls and the mean of genotyped test bulls is 0.424 units of SD of EBVs. This corresponds to fraction of 25% of animals not being genotyped. If the true R^2 from model [1] equals 0.65, the expected value of b_1 = 0.77 and the observed R^2 becomes 0.50.