

Interbull trend validation procedure

Principle

Before accepting data from a country to include in the international genetic evaluation, three trend validation tests are implemented. Data should pass all three tests, when applicable, in order to be included in the genetic evaluation of Interbull. Validation tests must be implemented when a country first enters the Interbull evaluation system and, thereafter, every time the country's national genetic evaluation system is modified, or when the last validation was more than two years ago.

Description

Development of the trend validation methods is credited to the scientific group of INRA in France. The following peer-reviewed publication provides deeper insight into the scientific principles on which the 3 trend validation tests are based:

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

Weller, J. I., Emanuelson, U. and Ephraim, E. 2003. Validation of genetic evaluation methodology using the nonparametric Bootstrap Method. *Interbull Bulletin* 31, 26-29.

Lidauer, M., Mäntysaari, E.A., Pedersen, J. and Strandén. 2005. Model validation using individual daughter deviations – statistical power. *Interbull Bulletin* 33, 195-199.

A brief description of the three testing methods follows:

Testing Method 1

Definition: Comparison of genetic trends estimated using only first lactation versus all lactations in the routine national genetic evaluations.

Responsibility: Countries using multiple-lactation models in national genetic evaluations are required to apply this method prior to delivering data to the Interbull Centre.

Motivation: Investigate the impact of cow records from different age groups on the genetic trend.

Data: Data used in the routine national genetic evaluation, whose results will be included in the international evaluation. The same time edit as applied in the genetic evaluation of Interbull should therefore be used.

Action: Estimate the genetic trend in the AI bull population with the usual model (b_T).

Repeat computations considering only first lactation records (b_1).

Compute the difference: $\delta = b_T - b_1$, for each trait.

- Criterion:** $|\delta|$ must be less than $.02 * \sigma_g$ or $.01 * \sigma_g$ if breeding values or transmitting abilities were considered in the trend estimation, respectively; σ_g is the genetic standard deviation for the trait; when each lactation is treated as different trait, instead of σ_g consider $(\sigma_{g1} * \sigma_{gT})^{**}.5 / r_{gIT}$, where σ_{g1} and σ_{gT} are the genetic standard deviations for first and all lactations (pooled) and r_{gIT} is the genetic correlation between first and all lactations (pooled).
- Remarks:** If $b_1 < b_T$ the genetic trend is over-estimated (the opposite if $b_1 > b_T$). Possible solutions: reconsider age pre-adjustment factors, include an age by parity or age within parity effect in the model, and include parity in the definition of management (contemporary) groups.

Testing Method 2

- Definition:** Analysis of within bull yearly Daughter Deviations (e.g. Daughter Yield Deviations, DYD), hereafter referred to as DD.
- Responsibility:** Countries where DD are available are required to apply this method prior to delivering data to the Interbull Centre.
- Motivation:** DD are independent of the year of calving of bulls' daughters. This method investigates the non-genetic time trend over the entire period considered in the national evaluation; deviations from zero will indicate biases in the genetic trend estimation.
- Data:** Within year DD computed in the most recent routine national genetic evaluation, whose results will be included in the international evaluations. All AI bulls first sampled in each country with daughters in at least 10 herds should be considered. For a test-day model individual daughter deviations are the sum of the additive genetic effect(s) of the sire, the Mendelian sampling term(s), the permanent environmental effect(s) pertaining to the daughter, and the residual term. A detailed description of computation of individual daughter deviations and (average) daughter deviations for test-day models is in Lidauer et al. (2005).
- Action:** Analyze DD with the following fixed effect model:

$$Y_{ij} = \text{BULL}_i + b * j + e_{ij} \quad [1]$$

where Y_{ij} is the DD considering daughters born in the j^{th} year of the i^{th} bull; by definition $j=0$ for the first year when at least 10 daughters of a bull were born; BULL is the effect of the i^{th} bull.

- Criterion:** $|b|$ from Model [1] must be less than $.01 * \sigma_g$, where σ_g is the genetic standard deviation for the trait.

Remarks: If $b > 0$ the environmental trend is under-estimated and the genetic trend is over-estimated (the opposite if $b < 0$)

Testing Method 3

Definition: Analysis of official national predicted genetic merit variation across evaluation runs.

Responsibility: Countries that can not implement Method 1 and 2 due to model restrictions are required to apply this method prior to delivering data to the Interbull Centre. Countries that have applied both Methods 1 and 2, may also apply this method if they wish, but it is not required.

The Interbull Centre implements this method with data from ALL participating organisations. Each participating organisation is responsible for delivering appropriate data to the Interbull Centre.

Motivation: Investigate the random variation associated with new daughters. The method is independent of country and can be applied at the venue of the international evaluations.

Data: Let YYYY be the year of the most recent national genetic evaluation. Data include bulls born between certain years (see section Time windows below), AI first sampled in each country, whose national genetic evaluation of year YYYY-4 was based only on first crop daughters (minimum 20 daughters in 10 herds). For each one of these bulls the following variables are needed:

X = National evaluation (predicted genetic merit) from the new model using the same data as in the last routine genetic evaluation run in year YYYY-4;

Y = National evaluation (predicted genetic merit) from the new model using the data from the most recent routine genetic evaluation run, whose results will be included in the international evaluation (year YYYY);

n_x = Total number of daughters on which X was based in year YYYY-4;

n_y = Total number of daughters on which Y was based in year YYYY;

n_j = Number of new (first calving) daughters considered in the last national evaluation of the j^{th} year, where $j=1$ for YYYY-3, $j=2$ for YYYY-2, $j=3$ for YYYY-1 and $j=4$ for YYYY;

t = Proportion of bull's new daughters, where

$$t = \sum_j \left(\frac{n_j}{n_y} * m_j \right)$$

where $m_j = (YYYY-4) + j$ - (mean year of first calving of daughters on which the bull's national evaluation in year YYYY-4 was based).

Time window: Appropriate time windows (birth years of bulls) may vary depending of the trait to be validated, the speed of their progeny test program and other

factors. A shift of the time window with one year will give a different set of bulls that qualify for the test. It is desirable to choose a time window so as to have the largest number of qualifying bulls as possible. It is required that bulls do have only first crop progeny for the evaluation in (YYYY - 4), and first and second crop progeny for the evaluation in YYYY. To fulfill these requirements for calving traits birth years commonly chosen are between YYYY-10 and YYYY-7 for direct calving ease / stillbirth), between YYYY-12 and YYYY-9 for maternal calving ease / maternal stillbirth and between YYYY-14 and YYYY-10 for production traits.

Action: Analyse Y for the i^{th} bull with the following model:

$$Y_i = a + b \cdot X_i + \delta \cdot t_i + e_i \quad [2]$$

weighing each observation by w
w=Weight of Y, where

$$w = \frac{(n_Y + k)^2}{n_X \cdot (n_X + k) \cdot (1 - R^2) + (n_Y - n_X)^2 \cdot \left(\frac{k}{n_Y - n_X} + \frac{k}{n_X + k} \right)}$$

where: k=variance ratio or $k = \frac{4 - h^2}{h^2}$, h^2 =heritability of the trait

R=expected correlation between methods used for calculation of Y and X (R = .99); h^2 should be provided by each country.

Precision: The empirical 95% confidence interval of δ should be submitted with the validation results. The empirical confidence interval can be obtained by a non-parametric bootstrap (Weller et al, 2003):

Repeat at least 1000 times:

- i. Sample (with replacement) among qualifying bulls. Sample size is the same as the number of qualifying bulls;
- ii. Compute the regression coefficient (δ) for this sample.

The 95% empirical confidence interval is obtained as the 2.5% and 97.5% percentile values of δ computed under ii.

Criterion: If δ from Model [2] differs significantly from zero (zero is not within the empirical 95% confidence interval) then $|\delta|$ from Model [2] must be less than $.02 \cdot \sigma_g$ or $.01 \cdot \sigma_g$ if breeding values or transmitting abilities were considered in the trend estimation, respectively, where σ_g is the genetic standard deviation for the trait.

Remarks: If $\delta > 0$ the genetic trend is over-estimated (the opposite if $\delta < 0$).

Logistics

The three testing methods must be implemented before a country sends data for an Interbull test-run. Results must be made available to the Interbull Centre on specifically designed forms that each country receives upon subscribing to the Interbull service.

Data for testing method 3 are sent from each country to the Interbull Centre according to a particular format that is described in the service package that each country also receives upon subscribing to the Interbull service.

When data from a country fail a test

Possible reasons for failing a validation test are 1) problems with the national genetic evaluation model in a country; 2) sampling error; 3) test is not appropriate for the genetic model in question.

The Interbull Centre investigates results in order to establish their statistical significance. With small data sets, discrepancies may sometimes not be statistically significant. In such cases, reasons for failing the validation tests are attributable to chance.

When results fail a validation test, and deviations from the established criteria are statistically significant, it is assumed that something is not completely correct with the national genetic evaluation model. Interbull Centre is then contacting this country for an effort to resolve the problem. In such cases, the aim is to revise the national genetic evaluation model so that results comply with the validation tests.