



Model Validation for Beef evaluations

EURC – Report

Years 2023 – 2024.

Background

Genetic evaluations estimate an animal's breeding value using information from the animal itself or its relatives. The reliability of these estimates depends on both the quality of the data collection systems and the methods applied to perform the estimation. The difference between an animal's true breeding value (TBV) and its estimated breeding value (EBV) is referred to as bias. Bias can significantly impact a population, as it may result in incorrect selection decisions. Therefore, it is crucial to assess whether the methods employed in an evaluation yield unbiased results. Validation methods have been developed to measure the extent of bias in evaluations.

These methods are particularly relevant in international evaluations, where data from multiple countries are integrated into a multi-country evaluation, each employing different models. In such cases, bias in one country's model can compromise the reliability of results in the international evaluation.

Currently, the situation of validation methods differs between dairy and beef evaluations. Interbull's international dairy evaluations have implemented validation methodologies since the mid-1990s, which have been continuously updated to align with advancements in evaluation methods. In contrast, the situation for beef evaluations is less developed. International beef evaluations (Interbeef) were launched for the first time in the mid-2010s, making it a relatively young initiative. Nonetheless, the need to establish robust validation methods for the countries participating in Interbeef evaluations has been recognised. To address this, a dedicated working group has been formed and has been actively working over the past few years to develop appropriate validation methods for beef evaluations. These methods must be validated and made available to the entire beef cattle breeding community across Europe.

Although Interbull methods are well established, Interbeef evaluations, along with beef evaluations in general, have specific characteristics that make validation distinct from the processes used in the international dairy context, as outlined below:

- Genetic evaluations are based on phenotypes linked to population-specific models.
- Maternal effects (both genetic and permanent environmental) are considered in most models.
- The level of connectedness between populations involved in the evaluations is lower, primarily due to the reduced use of artificial insemination (AI).



- Genetic trends are generally lower.
- Participating countries provide the models (which may differ slightly from those used at the national level) to be implemented in the international evaluation, along with the associated parameters (e.g., genetic variances, random effect variances, and direct-maternal co-variances).
- Genetic co-variances between populations are estimated as part of the Interbeef evaluation process.
- In the future, some countries may wish to join Interbeef without conducting a national genetic evaluation.

Given these factors, adapting or developing methods specifically for Interbeef evaluations—and beef evaluations in general—is a critical task. Recognizing the importance of this subject for the Interbeef WG, methodologies are studied together with the Interbull Centre and its members. Organizations from Ireland, Italy, and Sweden actively participate in the Beef Validation Methods WG.

Activities 2023 and 2024

From the outset, the Validation WG prioritised adapting existing Interbull methods to the beef evaluation context. The first method tested was the Interbull Method IV, also known as the Mendelian Sampling test. However, its results were inconclusive, leading the Scientific Advisors to recommend further investigation into Interbull Method II. Throughout 2023 and 2024, efforts have been dedicated to adapting this method for beef evaluations.

Interbull Method II

Originally developed for dairy evaluations, Interbull's Method II analyzes the within-bull variation of daughter yield deviation (Boichard et al., 1995). In the context of beef evaluations, we use the term *progeny deviation (PD)* since trait measurements are not limited to daughters. A PD is calculated for each observation, adjusting for the dam's breeding value and all model-included effects—except for the progeny's breeding value. This value is expected to be independent of environmental influences, allowing for an analysis of whether factors such as the year have any impact on the PD.

The model to analyse the individual deviations is shown as follow:

$$PD_{ij} = s_i + bj + e_{ij}$$

where PD_{ij} is the progeny yield deviation from sire s_i obtained in the year j ; j is the year of birth; b is the regression coefficient of year j ; and e_{ij} is the error. The year $j = 0$ is the year of birth of the first progenies of the bull.

We have to focus on the value of b , as it will indicate a within-bull yearly trend. If b is different from 0, then there is an environmental trend that is not considered by the model, and the genetic trend can be over ($b > 0$) or under ($b < 0$) estimated.

As a simple criterion to determine the pass or fail of the test, the absolute value of b ($|b|$) can not be higher than 1% of the genetic standard deviation ($\sigma_g * 0.01$).



Datasets

Data from three sources was used:

- Italy (ITA) Limousin (LIM) and Charolais (CHA) weaning weight evaluation
- Ireland (IRL) Limousin, Charolais and Simmental (SIM) weaning weight evaluation
- Limousin and Charolais data and models from the weaning weight Interbeef (ITB) international evaluation

The populations included in the data from Interbeef evaluations were:

- Limousin: Australia (AUS), France (FRA), Latvia (LVA), Denmark, Finland and Sweden (DFS), Germany (DEU), Estonia (EST), Czech Republic (CZE), Italy (ITA), Great Britain (GBR), Ireland (IRL) and Slovenia (SVN).
- Charolais: FRA, DFS, IRL, CZE, DEU, LVA, ITA, SVN and EST.

In order to compute Progeny Yield Deviations (PD), the models can not consider maternal effects; therefore, all the evaluations estimate only the direct effect of the trait. The PDs are calculated for each observation, subtracting all effects in the model plus half of the additive effect of the dam.

Regarding IRL and ITB, the PDs were obtained using Mix99. In the case of ITA, the evaluation was performed using BlupF90+ and the PDs were obtained using predictf90.

Criteria to include individuals in the test

The primary challenge in applying this methodology was determining the criteria for selecting bulls for the test, as the selection criteria used in dairy evaluations could not be directly applied to beef evaluations.

To identify informative bulls for the test, the group established several selection criteria:

- Progeny across at least three years (consecutive or non-consecutive)
- Number of progeny in the first year
- Number of progeny in subsequent years
- Number of herds

Different values for each criterion resulted in 6 scenarios.

Bulls were selected for each scenario and dataset (representing different countries), and the number of resulting bulls in each scenario was evaluated, along with whether the countries met the test criteria.

Additionally, a sensitivity test was conducted to assess the responsiveness of Method II across different scenarios. Three non-genetic trends were simulated within the datasets of various countries, corresponding to 0.5%, 1%, and 5% of the genetic standard deviation. The results were then analysed for each scenario to evaluate the method's performance.



After analysing all the results, the most promising and recommended scenario was the one where bulls were selected based on having at least three progenies per year, in at least three herds per year, across a minimum of three consecutive years. These criteria were then recommended for including bulls to validate models using Interbull Method II in beef evaluations.

These findings were presented to the Scientific Advisory Committee (SAC), who provided positive feedback, though some details required further revision.

Despite its potential to validate models in beef evaluations, the method still has limitations. For example, it cannot be used with models that include maternal effects, which are common in this type of evaluation. Consequently, the Validation WG plans to continue exploring other, more suitable methods, such as the LR validation method (Legarra and Reverter, 2018).

Status of the project

Currently, the Validation WG is reviewing the observations made by the SAC.

Since these observations are minor, guidelines are being developed for the application of Interbull Method II within the framework of beef evaluations.

Next steps

Although the results have been approved, we aim to expand the tests using Interbull Method II to include other breeds and traits, such as calving and carcass traits, in the second half of 2025.

Additionally, there are plans to begin testing the LR method within the framework of the international Interbeef evaluations.

References

Boichard, D., B. Bonaiti, A. Barbat, and S. Mattalia. 1995. Three Methods to Validate the Estimation of Genetic Trend for Dairy Cattle. *Journal of Dairy Science* 78:431–437. doi:10.3168/jds.S0022-0302(95)76652-8.

Legarra, A., and A. Reverter. 2018. Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genetics Selection Evolution* 50:1–18. doi:10.1186/s12711-018-0426-6.