ICAR INTERBEEF-WG

Beef Validation WG

Report May 19 2024

Members

Ross Evans, chairman (ICBF), Thierry Pabiou (ICBF), Stefano Biffani (CNR), Fernando Macedo (ITBC).

Background

Genetic evaluations estimate an animal's breeding value based on its or relatives' information. The reliability of this value depends on both the collection of information systems and the methods used to make the estimate. The difference between the *true breeding value* (TBV) and the *estimated breeding value* (EBV) of the animals under evaluation is called bias. Bias estimates can majorly impact the population as they can lead to wrong selection decisions. It is, therefore, important to determine whether the methods used in an evaluation obtain unbiased results. Validation methods have been developed to determine bias in an evaluation.

Validation methods have a particular interest in international evaluations, where information from different countries is integrated into a multi-country evaluation, in which each country has different models. In this case, a biased model from one country may lead to less reliable results in the international evaluation.

Nowadays, the situation concerning model validation is different for dairy and beef evaluations. The Interbull international dairy evaluations have had validation methodologies in place since the mid-1990s, and they have been constantly updated based on changes in evaluation methods. The situation at the beef evaluation level is different. Interbeef started with official evaluations in the mid-2010s, so it is still a relatively young service. However, the need to establish validation methods for the countries participating in the Interbeef evaluations has been identified, and a working group has been formed. Over the last few years, it has been working to develop validation methods for beef.

Although Interbull methods are well established, the Interbeef evaluations have specific characteristics which make validation different from the one performed in the dairy context, as described below:

- Genetic evaluations are based on phenotypes associated with population-specific models
- Maternal effects (genetic and permanent environment) are taken into account in most of the models, and maternal EBV is distributed
- Connectedness between populations involved in the evaluations is lower (because of the reduced use of AI.)
- Genetic trends are generally lower

- Participating countries provide the models (which could slightly differ from the ones used at national level) to be implemented in the International evaluation and the associated parameters (genetic variances, random effects variances, direct-maternal covariances intra or inter traits).
- The genetic covariances between populations are estimated within the Interbeef evaluation process instead.
- In the future, countries might be interested in joining Interbeef without a National genetic evaluation.

Considering the reasons above, adapting or developing methods for Interbeef evaluations is an important task.

Objectives of the Validation WG

The WG aims to develop methods to test and validate a set of population-specific model- parameters-performance, with a particular focus on the model used. According to several criteria (size of the population, genetic diversity of the breed), the validation test result aims to provide a "yes/no" answer.

The outcomes from this WG are very important for two main reasons:

1- First, the validation of the data from the countries participating in the Interbeef evaluation can serve as a criterion for their inclusion or non-inclusion and thus improve the accuracy of the evaluations.

2- A second important application is to establish these methods as standards for validating country models at the national level. In this way, countries can have a tool to demonstrate that their evaluations are unbiased.

Validation WG Activities 2023

During 2022, the WG worked on adapting the Method IV (Mendelian sampling) of Interbull to beef. However, the results were not conclusive, and on the other hand, given the progress in genomic evaluations, it was discussed with the scientific advisors that the method IV might not be the most suitable methodology for beef evaluations. At that moment, Method II was identified as possibly being used in the Interbeef framework.

Therefore, during 2023, the WG focused on adapting Method II of Interbull to beef evaluations.

Method II of Interbull analyses the within bull progeny deviations, investigating the non-genetic time trend over the years. In the case of dairy cattle, the daughter yield deviation is used; in beef, we investigate the progeny deviation.

The WG first identified the criteria to include individuals in the evaluation and then validated the method.

The main results of this work are presented in this report.

Material and methods

Datasets

We used data from three sources:

- 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 a Mix99 (MiX99 Development Team, 2022). In the case of ITA, the evaluation was performed using BlupF90+ and the PDs were obtained using predictf90.

Methods

Method II of Interbull

Originally defined for dairy evaluations, Method II of Interbull analyses the within-bull variation of daughter yield deviation; in a beef framework, the variation of progeny yield deviation. A PD is calculated for each observation, adjusted by the dam breeding value and the effects included in the model, except for the progeny breeding value. It is expected that this value is independent of any environmental effect (Boichard et al., 1995), so it is possible to analyse if the effect of the year has some 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 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 ($\sigma_g * 0.01$).

Criteria to include individuals in the test

In order to select informative bulls to participate in the test, the group worked on the different criteria:

- Progeny in at least three years
    - The years may be consecutive or not
- Number of progeny in the first year
- Number of progeny in the following years
- Number of herds

Different values in each criterion led to several scenarios.

The WG decided on six final scenarios after different approaches and based on the number of bulls included in the test. All the scenarios consider at least three consecutive years. They differ in the number of progeny in the first year, the number of progeny per year and the number of different herds in total.

The final scenarios tested by the WG were defined as below:

A1 – at least five progeny per year and at least three different herds in total

A2 – at least five progeny per year and at least five different herds in total

B1 – at least five progeny in the year 0, at least two progeny in the following years and at least three different herds in total

B2 – at least five progeny in the year 0, at least two progeny in the following years and at least five different herds in total

C1 – at least three progeny per year and at least three different herds in total

C2 – at least five progeny in the year 0, at least three progenies in the following years and at least three different herds in total

Even when test was done using data from ITA and IRL and for CHA and LIM breedsin his report, we will only include the results for the above scenarios using data from the International evaluations and LIM animal weaning weight.

Executive support: **Interbull Centre**
Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences - SLU
Ulls väg 26, PO Box 7023, 750 07 Uppsala, Sweden
E-mail: interbeef@slu.se, URL: www.interbull.org

ICAR INTERBEEF-WG

Sensitivity of method II in beef context

The second step of the research focused on analysing the sensitivity of the Interbull Method II for beef evaluations. Method II analyses the presence of a yearly trend of environmental origin not captured by the evaluation model.

Two ways of performing this test were discussed. We could rerun the evaluation by omitting some effect, introducing an environmental trend, or introducing noise in the PD values, thus generating an artificial yearly trend.

The group opted for the second option, as applying different noise types and values would be easier.

Noise was introduced in different ways:

- By bull
- Per year
- By bull and year

This report will include the results when the noise was introduced by year.

Results

Below, we present the main results of the investigation for LIM animal weaning weigth data from International evaluations

For all countries, the entire period of the evaluation was considered. The exception was FRA, only the last 15 years of data were considered for ease of the computation of the results.

Application of method II in beef evaluations

The first step in this study was to test Method II based on different scenarios.

A summary of the main results is presented below. The number of bulls participating in the tests was included in brackets for each country.

Table 1 present the results for Limousin animal weaning weight data.

Table 1. Results for Limousin aww data.

Country	Sc. A1	Sc. A2	Sc. B1	Sc. B2	Sc. C1	Sc. C2
AUS	241	185	334	264	338	290
CZE	115	67	233	175	195	178
DEU	362	181	549	353	534	458
DFS	877	537	1012	689	1192	973
EST	11	--	27	14	18	17
FRA	1936	556	1976	640	2199	1937
GBR	442	315	598	440	621	528
IRL	182	154	282	233	290	228
ITA	196	64	336	183	300	267
LVA	22	8	39	22	35	28
SVN	9	--	24	18	15	14

Numbers = number of bulls included in the test. Green = Passed. Red = failed

The most restrictive scenarios were A1 and A2, as at least five progeny per year were required for each sire. On the other hand, scenarios B1 and B2, which considered at least 5 progeny in year 0 and then at least 2 in the following years, ended up being less restrictive than scenarios A1 and A2, achieving a higher number of bulls participating in the test. The scenario that achieved the highest number of participating bulls was the C1 scenario, where at least three progeny per year and at least three herds in total were required. Countries with smaller populations (EST, LVA and SVN), in general, did not achieve more than 100 bulls included in the test and in most cases, this number was less than 50. These cases, even if they could pass the test, were not considered for the sensitivity trial, as tests with too few bulls can lead to biased results.

Regarding the differences in results across scenarios, when the number of bulls included was greater than 100, the results were relatively consistent. The exception was for the A2 scenario which presented some different results than the others. For example, FRA in the LIM breed passed the test except for the A2 scenario, the most restrictive, which included 556 bulls. Figures 1, 2 and 3 show the annual trend for FRA in the A2, B1 and C1 scenarios.Thus, we observe that the selection of bulls based on different criteria can affect the test result, mainly for those countries that may have a trend of an environmental effect not contemplated in the model with a slope close to 1% of σ_g. Very restrictive scenarios may lead to bias in the application of the method.

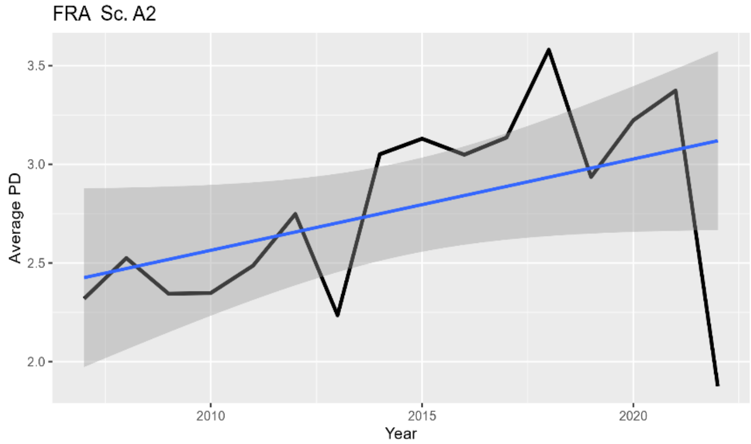

Figure 1 – PD trend for France in Scenario A2

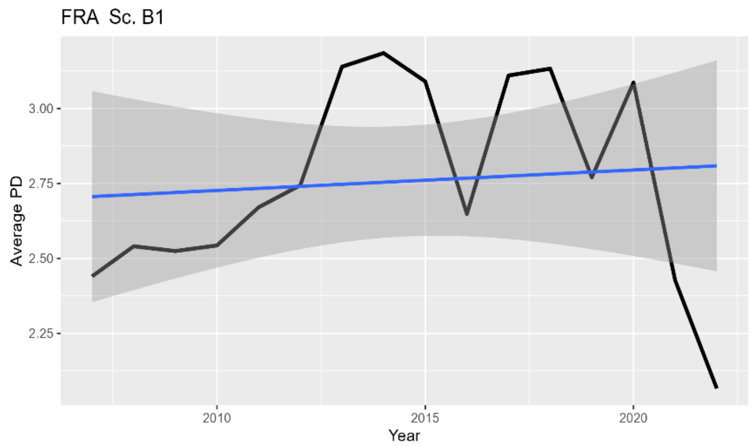

Figure 2 – PD trend for France in Scenario B1

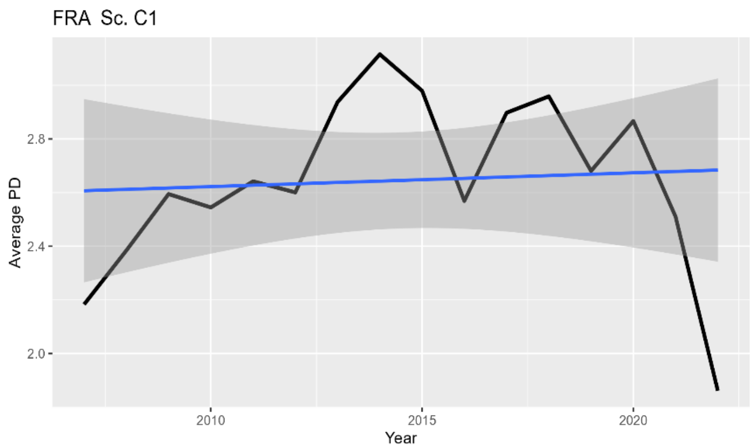

Figure 3– PD trend for France in Scenario C1

Sensitivity of Method II

Tables 2 show the results obtained for each scenario where yearly noise was introduced, representing 0.5%, 1% and 5% of σ_g. Only those scenarios where countries passed the test with a number of bulls greater than 100 were studied.

Table 2. Results for Limousin aww senitivity trial.

Country	Noise	Sc. A1	Sc. A2	Sc. B1	Sc. B2	Sc. C1	Sc. C2
AUS	0.5%	✓	✓	✓	✓	✓	✓
	1%	X	X	X	X	X	X
	5%	X	X	X	X	X	X
CZE	0.5%	✓	--	✓	✓	✓	✓
	1%	✓	--	✓	✓	✓	✓
	5%	X	--	X	X	X	X
DEU	0.5%	X	--	✓	X	✓	✓
	1%	X	--	X	X	X	X
	5%	X	--	X	X	X	X
DFS	0.5%	✓	✓	X	X	✓	✓
	1%	X	X	X	X	X	X
	5%	X	X	X	X	X	X
FRA	0.5%	✓	--	✓	X	✓	✓
	1%	X	--	X	X	X	X
	5%	X	--	X	X	X	X
GBR	0.5%	✓	✓	✓	✓	✓	✓
	1%	✓	✓	X	X	✓	✓
	5%	X	X	X	X	X	X
IRL	0.5%	✓	✓	✓	✓	X	✓
	1%	X	X	✓	X	X	X
	5%	X	X	X	X	X	X
ITA	0.5%	--	--	--	✓	--	--
	1%	--	--	--	✓	--	--
	5%	--	--	--	X	--	--

The repetition of the tests, including simulated trends, sought to test the sensitivity of Method II. It would be expected that if countries pass the test because there is no trend associated with the year of birth of the progeny, they would do so until a simulated trend of 1% of σ_g is added.

This was the case for many of the countries. However, we also observed the situations where countries failed the test when only 0.5% of noise was added and in some other cases even with 1% of noise, the countries could pass the test. The first situation occurred when the test of the data without noise was passed, but the slope (b) value was very close to the limit (1% σ_g). Thus, when even a small amount of noise (0.5%) is added, this trend ultimately turns out to be greater than $\sigma_g * 0.01$.

Moreover, when noise corresponding to 1% of σ_g was added, and the result was still positive, it was because the original trend was negative. To pass the test, the absolute value of b cannot be greater than 1% σ_g. However, b can be negative, and the addition of noise was positive, so this addition of 1% noise transformed a negative trend into a positive one but still below the limit. For example, for GBR in the C1 scenario (Figure 4), as a criterion to pass the test, $|b|$ should not exceed 0.1667 (> to 1%σ_g). The test without noise obtained a trend of -0.0399; after adding 0.5% noise, the trend changed to 0.0435 and when 1% was added, b= 0.1269. In none of the cases b was greater than the limit. This value was only exceeded when 5% noise was added (b=0.7937).

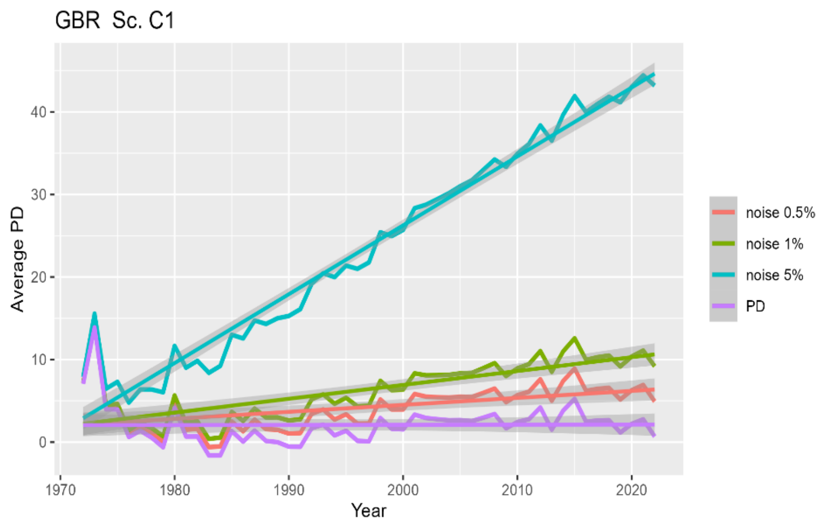

Figure 4. Yearly PD trend for LIM GBR in Sc. C1 with different addition of noise.

General conclusions

In conclusion, we can summarise:

- Regarding the selection of the criteria, given that the results across scenarios were similar and that the higher the number of bulls included in the test the better the reliability of the test, the use of at least three consecutive years, with at least three progeny per year and at least three different herds in total (scenario C1) is advised as a criterion.
- The test proved to be useful for detecting yearly trends of non-genetic origin when using the progeny deviation per year.

Final remarks

Tests have also been performed on other breeds, but the results are still being reviewed. After their analysis, the results will be included in this report, and an updated version will be sent to the participants of the Interbeef Technical Committee.

This report will serve as a basis for submission to the SAC members for approval of the Interbull Methode II in the framework of beef genetic evaluations.

However, given the limitations of the test, the WG will continue to investigate other validation methods.

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.

MiX99 Development Team (2022). MiX99: A software package for solving large mixed model equations. Release I/2022. Natural Resources Institute Finland (Luke). Jokioinen, Finland. URL: http://www.luke.fi/mix99