

Benefits of Performing Interbull Validation on National Models

Purpose of Genetic/Genomic Evaluations and Their Economic Values to the Farmers

The cattle breeding industry has reached very high level of production both in matters of quantity (of milk or meat produced) but also, and most importantly, in the organoleptic characteristics of such products. Both are very important aspects in order to provide high quality nutrients to an ever growing world-wide human population. This important achievement could not have been reached without a continuous work of genetic and, more recently, genomic selection world-wide.

Starting with genetic selection, all animals' performances together with their pedigree relationships have been recorded and included in even more sophisticated statistical models in order to better distinguish which amount of the animal's performance was due to genetic and, therefore, inheritable factors and which, instead, was due to the particular environmental conditions the animal lived in. This is a very important distinction that still allows national genetic centres and farmers alike, to better estimate the amount of desirable genes each animal would be able to transmit to its offspring and, by doing so, estimating the genetic worth (i.e. breeding value) of such animals and make educated decisions on which animals to retain for breeding.

In the last decade, a new type of selection, called genomic selection, based on the analysis of the animal's DNA, rather than on its performances, has been implemented in many countries mostly thank to three specific factors: 1) The considerable reduction in the generation interval required to estimate the genetic worth of any young animals: as the animal's performances are not required for this type of evaluation, which is based on the direct analysis of the animal's DNA, the genetic worth of any animal can be estimated at a very early stage of the animal's life. This has the direct benefit to speed up selection while reducing the costs of keeping undesired animals in the farm. 2) The substantial price reduction required nowadays for analyzing (genotyping) animals: the continuous improvements at the laboratory level in the techniques applied for the analysis and in the time required, has allowed for a significant price reduction moving from several hundred dollars to just a few dollars per animal's genotype, making genotyping more cost-effective and accessible to a wider range of farmers. 3) The availability of faster computing machines: The amount of data to process and the new statistical models required did increase exponentially in matter of size and complexity so much so that the computational time, rather than the availability of data, was often the limiting factors for applying this specific type of selection. Nowadays, the fast development in the computing industry is overcoming this particular problem. 4) The fast spread-out of new reproduction technologies, as sexing semen, has strongly increased the use of young TOP genomic evaluated bulls. By using sexed semen the cattle farmers can reduce the number of young animal for replacement with benefits for economic and environmental sustainability of cattle farms.

The two sets of evaluations, applied separately or combined, have the ultimate goal to help farmers managing their herds while improving their productivity, general health and welfare level, environmental sustainability and, therefore, also the farmers' revenue.

Genetic bias and its consequences for the evaluation

While estimating the genetic worth (i.e. breeding value) of the animals within a herd is an important tool allowing the farmers improving managing the farm by reducing susceptibility to diseases, improving fertility and general wellness of the animals while maintaining the production level stable, or even increasing it, the assessment of an animal's genetic worth would still be based on a statistical estimate as, unfortunately, the true genetic worth would still remain unknown. As any other statistical estimates, even the breeding value of an animal is subjected to a statistical bias that would lead to an over (or under) estimation of the animal's genetic worth. Bias can arise from incomplete data collection, over/under parameterized of the statistical models used or not enough records available. If left unhandled, the initial bias would tend to accumulate in all subsequent evaluations leading to unrealistic estimates that could penalized (negative bias) or overestimate (positive bias) an animals' breeding value. Dealing with biased estimations would, in the long run, jeopardized the outcome of the genetic selection, as farmers would be recommended of focusing more on the overestimated animals which would, in turn, not be able to performed according to their estimation which ultimately would lead to an economical loss for the farmer. Bias estimations are, therefore, always an issue, even more so when the selection is applied on a national or international bases.

Moreover, the large use of genomic preselected young bulls can affect the general reliability of the genomic evaluation by introducing a common bias called “Genomic Pre Selection Bias” (GPS). Also for this reason, specific validations techniques are needed and they can be finetuned just at international level i.e. based on international comparisons.

The purpose of Validation

Although the bias can be considered as intrinsic in any type of statistical estimations, a set of checks can be put in place in order to quantify its magnitude. The set of checks can vary depending on the type of data and models at hand. In the animal breeding industry, Interbull has over the years developed several approaches to assess the bias level in the different national genetic centres’ estimates. Such approaches are known with the name of Validation methods and are applied by Interbull Centre to assure that the data received by countries would be as unbiased as possible and, when detected, the bias would never exceed a tolerated threshold of 2%. Interbull’s validation methods have a twofold purpose:

- 1- To avoid that highly biased data could enter the genetic/genomic international evaluations that could jeopardize the quality of the international estimated breeding values.
- 2- To provide feedback to member organisations on which statistical model(s) would require a closer revision because causing results to have a bias above the tolerated threshold.

For this reason, Interbull does require member organisations to provide validation results for all their breeds and traits any time that:

- A country joins the international evaluations for the first time
- The statistical model is new or substantially modified
- The data used to feed the statistical model has changed (both in nature or in its size)

In the event that neither of the three cases described above would be applied, Interbull would still require validation results if the last available are older than two years.

Interbull validation is applied to genetic and genomic national evaluations alike and with the same criteria as described above. The tight validation requirements do allow Interbull to perform a strict check on the quality of the data used as input to the international evaluations while for the different countries it provides a quality check on the methodologies applied and, eventually, information on which models would need to be checked/modified because leading to bias estimates.