SMARTER

SMAll RuminanTs breeding for Efficiency and Resilience

Research and Innovation action: H2020 – 772787
Call: H2020-SFS-2017-2
Type of action: Research and Innovation Action (RIA)
Work programme topic: SFS-15-2016-2017
Duration of the project: 01 November 2018 – 31 October 2022

Method for assessing potential bias due to design of cross validation analysis

Andres Legarra

INRAE

* Deliverable leader – Contact: andres.legarra@inrae.fr

DELIVERABLE D5.3

Workpackage N°5

Due date: M36

Actual date: 28/10/2021

Dissemination level: Public

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 772787

About the SMARTER research project

SMARTER will develop and deploy innovative strategies to improve Resilience and Efficiency (R&E) related traits in sheep and goats. SMARTER will find these strategies by: i) generating and validating novel R&E related traits at a phenotypic and genetic level ii) improving and developing new genome-based solutions and tools relevant for the data structure and size of small ruminant populations, iii) establishing new breeding and selection strategies for various breeds and environments that consider R&E traits.

SMARTER with help from stakeholders chose several key R&E traits including feed efficiency, health (resistance to disease, survival) and welfare. Experimental populations will be used to identify and dissect new predictors of these R&E traits and the trade-off between animal ability to overcome external challenges. SMARTER will estimate the underlying genetic and genomic variability governing these R&E related traits. This variability will be related to performance in different environments including genotype-by-environment interactions (conventional, agro-ecological and organic systems) in commercial populations. The outcome will be accurate genomic predictions for R&E traits in different environments across different breeds and populations. SMARTER will also create a new cooperative European and international initiative that will use genomic selection across countries. This initiative will make selection for R&E traits faster and more efficient. SMARTER will also characterize the phenotype and genome of traditional and underutilized breeds. Finally, SMARTER will propose new breeding strategies that utilise R&E traits and trade-offs and balance economic, social and environmental challenges.

The overall impact of the multi-actor SMARTER project will be ready-to-use effective and efficient tools to make small ruminant production resilient through improved profitability and efficiency.

1 Summary

Evaluating bias and accuracy in small ruminant genetics is not easy because males have little number of phenotyped offspring, because studies are generally small, and often use hghly selected (elite) animals. In this document we present some existing methods and we argue that they are either difficult to use or incorrect. Here we propose the general use of method LR ("Linear Regression"), with small examples, and discuss practicalities. During Smarter, we have applied the method to several data sets, gathered experience, and further developed some methodological aspects.

2 Introduction

In genomic selection, use of early genomic proofs can lead to suboptimal selection decisions if there is bias (see below for description of bias). In sheep and goats, and in particular for traits expressed in females, there is a lack of good tools to evaluate the presence or absence of bias, and the methods to evaluate accuracy of genomic selection are suboptimal.

We do we concern about bias? Selection theory establishes that selection is optimal if each candidate to selection is compared fairly to each other. This means that across individuals, Estimated Breeding Values (EBV, $\hat{u}$) of the selected candidates is equal to the expectation of the (true) Breeding Values (BV, u). When animals are selected, this is true under two conditions: $\bar{u} = \bar{\hat{u}}$ and $cov(u, \hat{u}) = var(\hat{u})$, where the means and the covariances apply across the animals selected in an operation (i.e. at the time of selecting young male lambs). The property $cov(u, \hat{u}) = var(\hat{u})$ is needed because if the distribution of $\hat{u}$ of e.g. young animals is too (or not enough) spread, we will select too many (or too little) young animals. Note that at this point, these properties are not statistical and therefore are neither "frequentist" nor "Bayesian".[1]

These properties can be formalized as

(1) *equality of estimated and true means* :

$$\mathbf{1}'\hat{\mathbf{u}} = \mathbf{1}'\mathbf{u}$$

or equivalently $\frac{1}{n}\sum\hat{u}_i = \frac{1}{n}\sum u_i$ or still $\bar{\hat{u}} = \bar{u}$, and

[1] The compensation that NZ farmers got in 2010 for using genomic bulls with biased genomic proofs did not know about priors or sampling distributions *:-)*

(2) *slope of true on estimated equal to 1*

$$\frac{1}{n}\sum\left(\hat{u}-\overline{\hat{u}_i}\right)^2=\frac{1}{n}\sum[\left(\hat{u_i}-\bar{\hat{u}}\right)(u_i-\bar{u})]$$

or equivalently $cov(u,\hat{u})=var(\hat{u})$.

Henderson (Henderson (1975), Henderson (1982)) established that the two properties above hold, even if there is selection, *on expectation* for *one* animal across repeated conceptual sampling of its $(u,\hat{u})$. Then Legarra and Reverter (2018) proved that the proof applies to *sets* of EBVs from groups of animals, so we have that the two properties hold *on expectation* for *many* animals across repeated conceptual sampling. By the Law of Large Numbers, when the number of animals is large, a number converges to its expectation. This means that, for a *large* number of animals, $\bar{\hat{u}}=\bar{u}$ *must* hold empirically.

So the theory says that, without invoking some esoteric statistical framework, genetic evaluations should be unbiased. But how can we check this? We don't have u, only $\hat{u}$. In dairy cattle, they compare predictions vs. progeny proofs (or Daughter Yield Deviations) but in other species the number of offspring of each animal is small.

In addition, we're interested in finding out the accuracy of genomic prediction, i.e. $r(u,\hat{u})$. Again, it is difficult to obtain this number in small ruminant cases.

3 Bias due to using pre-corrected data or De-Regressed Proofs (DRP)

The following is extracted from Legarra and Reverter (2018). Often we have used precorrected phenotypes y^* or deregressed proofs, and compare predictions $\hat{y}$ with (precorrected) observations y^* (this method is sometimes called "predictability"). The estimator of accuracy is e.g. $r\approx cor(y^*,\hat{y})/h$ for h^2 the heritability (Legarra et al. 2008). But this ignores that precorrection generates a covariance structure in y^*, is *very* sensitive to low values of h^2, and it also ignores that animals used in these studies can be preselected (case for instance of elite males). This leads to paradoxes:

- $r>1$ (observed in chicken)
- $r_{pedigree}>r_{genomic}$ (observed in dairy cattle for fertility)

3.1 Bias due to ignoring the effect of selection on genetic variance

It also ignores that candidates to selection have reduced genetic variance (Bijma (2012)). For instance, for prospective AI rams in dairy sheep, because they're highly selected, their genetic variance is less than the "normal" genetic variance.[2]

[2] Note that the genetic variance is recovered when this animals mate to females in the next generation.

Consider for instance that we made a study on growth in meat sheep in selected rams in a performance recording station. These rams are selected based on parent average and therefore their genetic variance, is, say, $k = 80\%$ of the populational one, and $h^2 = 0.3$. Through cross-validation we obtain $cor(y^*, \hat{y}) = 0.4$ and we conclude that $r \approx cor(y^*, \hat{y})/h = 0.73$. However this is incorrect because these animals were selected, so that *in these rams*, the heritability is actually $h^{2*} = \frac{kh^2}{1-(1-k)h^2} \approx 0.26$. Coupling in our equation $r \approx cor(y^*, \hat{y})/h^* = 0.78$, quite higher.

3.2 Bias due to pre-correction by fixed effects

There is a second, non-negligible source of bias. We use y^* (precorrected data) as it was "exact". This leads to overestimation of accuracies. In Legarra and Reverter (2018) we worked out that for a balanced design with n_i records per contemporary group, the bias is such that the *relative* overestimation of accuracy is of order $\frac{1}{n_i}$. For instance:

- Dairy sheep: assume 25 animals / contemporary group. This leads to overestimation of accuracy by $1/25 = 4\%$. If $r \approx cor(y^*, \hat{y})/h = 0.73$, this r was overestimated by 4% so that the actual accuracy should be $r = 0.73(1 - 0.04) = 0.70$.
- Beef cattle: 5 animals / contemporary group. This leads to overestimation of accuracy of 20%

4 Method LR to the rescue

For all these reasons we want better methods to assess biases and accuracies.

Legarra and Reverter (2018), with further proofs by Bermann et al. (2021), extended the machinery developed by Henderson (1975) and Reverter et al. (1994) to infer biases and accuracies by splitting the data set. They defined *partial* (*p*) and whole (*w*) data sets, so the *partial* data set contains all information until a given date and the *whole* data set contains all information available for the analyst until a later date (not necessarily now). The procedure, called LR from Linear Regression[3] is described next. LR in a nutshell

You have complete (*whole*) records, pedigree and (perhaps) markers. Consider a cut-off date. Records before these date make the *partial* data set: $\mathbf{y}_p$ whereas all records make the *whole* data set: $\mathbf{y}_w$. Then you run two genetic evaluation with either the *partial* data or the *whole* data, and you keep the entire pedigree and markers in both. In these manner, you have EBVs for all animals in both cases, $\hat{u}_p$ and $\hat{u}_w$ respectively.

[3] The fact that the initials of the authors are LR is, of course, coincidental *:-)*

Then you compare EBVs of animals in *partial* and *whole* prediction. You don't include *all* animals; you consider contemporary animals with similar information, in which you have an interest (for instance, males candidates to selection). We call this *focal* animals or *focal* groups. See below for examples.

The comparison is very simple and it just consist in a series of statistics that can be easily computed. We propose several criteria. This can be found in F. L. Macedo et al. (2020) which is the most up-to-date source. Note that in the following, whenever we put something like $cov(\hat{\mathbf{u}}_p, \hat{\mathbf{u}}_w)$ we mean a $scalar$ (the "observed" covariance) and not a *matrix* (which is the sampling or prior distribution of the vector).

4.1.1 Bias

This is measured using $\hat{\Delta}_p = \bar{\hat{\mathbf{u}}}_p - \bar{\hat{\mathbf{u}}}_w$. The expectation is 0 (no bias). A positive value means that animals with *partial* information are overevaluated.

4.1.2 Slope

Also calledr over/underdispersion. This is measured using $\hat{b}_p = \frac{cov(\hat{\mathbf{u}}_p, \hat{\mathbf{u}}_w)}{var(\hat{\mathbf{u}}_p)}$ or, equivalently, computing the slope b_1 of the linear regression "whole on partial" $\hat{u}_w \sim b_0 + b_1 \hat{u}_p + \epsilon$. The expectation is 1 (no over- neither under-dispersion), values lower than 1 mean that selected candidates are overestimated. This is the kind of bias commonly reported in dairy cattle studies.

A very small example with 5 individuals follows:

```
# these are actually 5 "proven" bulls
EBV2018=c(999,849,831,953,764)
EBV2019=c(973,833,904,963,807)
Delta_p=mean(EBV2018)-mean(EBV2019) # -16.8
b_p=cov(EBV2019,EBV2018)/var(EBV2018) #0.71
aa=lm(EBV2019~EBV2018)
b_p=aa$coefficients[2] # 0.71
```

4.1.3 Accuracies

There are two estimators of *relative* accuracies and two estimators of *absolute* accuracies.

- The first statistic is the correlation between *partial* and *whole* EBVs: $\hat{\rho}_{wp} = \frac{cov(\hat{\mathbf{u}}_p, \hat{\mathbf{u}}_w)}{\sqrt{var(\hat{\mathbf{u}}_p) var(\hat{\mathbf{u}}_p)}}$ (or simply $cor(u\_p, u\_w)$). This has expected value $\frac{acc_p}{acc_w}$ where acc means accuracy.

So, this estimates a *ratio* of accuracies and *not* the absolute accuracy. For instance, Values close to 1 indicate that "partial evaluation" was "as accurate" as "whole" evaluation, but both evaluations could be "little accurate".

A byproduct of $\hat{\rho}_{wp}$ is an estimator of the *relative increase in accuracy*. In effect, $\frac{1}{\hat{\rho}_{wp}} - 1$ has expected value $\frac{acc_w - acc_p}{acc_p}$, which is the relative increase in accuracy from *whole* to *partial* . For instance, boars can be evaluated for carcass traits *before* or *after* some full-sibs have been slaughtered, and $\frac{1}{\hat{\rho}_{wp}} - 1$ gives the relative increase in accuracy.

- The second statistic is $\hat{\rho}^2_{wp} = \frac{cov(\hat{\mathbf{u}}_p, \hat{\mathbf{u}}_w)}{var(\hat{\mathbf{u}}_w)}$, with expected value $\frac{acc_p^2}{acc_w^2}$, i.e. the ratio of *reliabilities* (squared accuracies).

Note that in fact this statistic $\hat{\rho}^2_{wp}$ is the slope b_1 of the regression "partial on whole": $\hat{u}_p \sim b_0 + b_1 \hat{u}_w + \epsilon$. A note of caution of this statistic is that the expected value requires that the evaluation is unbiased ($\hat{b}_p = 1$) something that is *not* required for $\hat{\rho}_{wp}$. In principle, the value obtained for $\hat{\rho}^2_{wp}$ should be the square of the value obtained for $\hat{\rho}_{wp}$, but this is not true in practice as it holds only in expectation.

Both statistics are easy to compute:

```
rho_pw=cor(EBV2018,EBV2019) # 0.9101622
rho2_pw=cov(EBV2019,EBV2018)/var(EBV2019)# 1.15944
```

note that in this example $\hat{\rho}^2_{wp}$ is not admissible ($\hat{\rho}^2_{wp} > 1$ would mean that $acc_p > acc_w$) and this is because in the example $\hat{b}_p$ is not even close to 1.

- The first estimator of *absolute* reliability is an estimator of "selected" reliability: $\widehat{acc}_p^2 = \frac{cov(\hat{\mathbf{u}}_p, \hat{\mathbf{u}}_w)}{\sigma^2_{u^*}}$. The denominator $\sigma^2_{u^*}$ is the variance of animals in the focal group (and not the variance of the base generation σ^2_u).

When animals are pre-selected (for instance, prospective AI rams selected based on parent average) their genetic variance $\sigma^2_{u^*}$ is less than the "normal" genetic variance σ^2_u. As an example, in Manech Tete Rousse, $\sigma^2_u \approx 500$ but $\sigma^2_{u^*} \approx 350$ for young selected rams (for milk yield) Macedo, Christensen, and Legarra (2021). The variance $\sigma^2_{u^*}$ can be estimated using Gibbs Sampling (Sorensen, Fernando, and Gianola (2001),F. L. Macedo et al. (2020)).

So, this equation gives the "selected" reliability (Bijma (2012),Dekkers (1992)), which is the "ability" to rank *within* those animals (more difficult when they are selected). However, we can't (easily) use this reliability to predict genetic progress, and we can't compare it with results in less selected animals, say, females. Also, the numbers do not match with those model-based, i.e. by Selection Index theory or from the inverse of the MME. The solution to this was given by Dekkers (1992) and Bijma (2012), and it leads to the last statistic:

- Unselected reliability, $\widehat{rel}_p = 1 - \frac{\sigma^2_{u^*}}{\sigma^2_u}(1 - \widehat{acc}_p^2)$. The mathematical explanation of all this is quite boring and convoluted, but some detailed exapmles can be found in F. Macedo, Reverter, and Legarra (2020) and F. L. Macedo et al. (2020).

4.2 Examples of interpretation

Just to give a feeling of what these numbers look like and mean. When we did the first cross-validation approaches in dairy sheep, we used AI rams that after selection based on parent average, were used in progeny testing. In order to compute if genomic selection is good, we can evaluate these rams with ssGBLUP at birth, and then after progeny. The first result that we get is $\hat{\rho}_{wp}$, but it can't be used to predict genetic progress of genomic selection. Then we do better and we compute $\widehat{acc}_p^2$, but we obtain a number that is very small because the animals are highly selected. What we want is the accuracy of the genomic young rams *if they were not selected*, because a genomic selection scheme genotypes a wide basis of animals. To do so, we use the equation above to transform $\widehat{acc}_p^2$ into $\widehat{rel}_p$, which is the number that we want.

For instance, we obtained the following Table (F. L. Macedo et al. (2020)):

Method	$\widehat{acc}_p^2$	$\widehat{rel}_p$	$\hat{\rho}_p^2 w$
BLUP-MF	0.22	0.53	0.32
SSGBLUP-MF	0.32	0.59	0.45

In the Table, the numbers of $\widehat{acc}_p^2$ seems "obviously wrong" because, for instance, for BLUP the reliability of the Parent Average from progeny-tested sire and phenotyped mother is usually close to 0.5, much higher than the observed numbers of ~ 0.25. However, the $\widehat{acc}_p^2 = 0.22$ in BLUP is the reliability *within the selected rams*, whereas the reliability *across all possible* rams is in fact $\widehat{rel}_p$, which has a value of 0.53 much closer to what we expect. The value of $\rho_p^2 w$ is more complicated to interpret. However, in the three columns it is obvious that SSGBLUP is more accurate than BLUP.

4.3 Practicalities

1. You evaluate the bias and accuracies for a category of animals. We call this *focal* animals or *focal* groups. These are contemporary animals for which the properties above hold, which are "exchangeable" (in other words, we're interested in the group, not in each individual animal) and in which we are interested. For instance young born rams can be a focal group. 1st-lambing females can be a focal group, and rams with first crop of daughters could be a focal group as well. But it is not a good idea to define a focal group composed of 50% progeny-tested rams 4 year old and 50% young animals that are 1 year old, because the first will be more accurate and the second more shrunken towards the mean. To define the focal animals, the best way is to do it by analyzing the data: for instance, take all m rams born in year say 2010, and from them select those n that had offspring with record in 2014, but not before. Then the number of animals in the focal group is n.

2. Define dates in a way such that the focal individual will have more information in the *whole* than in the *partial* data set. For instance, young rams could have only parents' (and genomic) information in the *partial* data set and offspring information in the *whole* data set. First lambing females could have one record for milk yield in the *partial* and two records in the *whole* ; and similar cases. In the example above, the year of *partial* can be 2010 and the year of *whole* can be 2014.

3. The way we do this is using the data set and "looking forward" from each year. For instance, we take all rams born in 2014 that were used in AI , and few years later (say 2017) we find out which of these rams have daughters with milk yield. This defines a focal group for "partial"=2014 and "whole"=2017 We can do the same for 2014 vs. 2018, 2019, etc.

4. In these manner we have many "pairs" of *whole* and *partial*. For instance you can do "partial" at 2010, 2011,… and compare each of them vs. "whole" at 2014, 2015… . It is important to do several comparisons because the statistics vary a lot across years. Using several pairs of *whole* and *partial* requires automatic handling of files and data editing, that we do using automated scripts in R, Unix tools, and R scripting. The genetic evaluations, themselves, can be run in any software that you like.

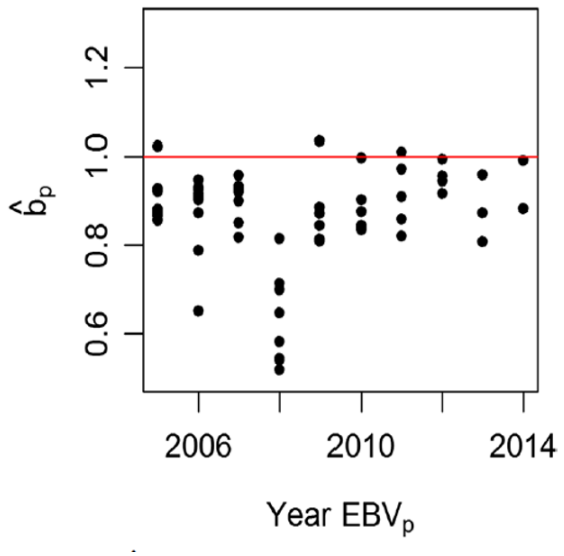
5. In practice we delete "records" (milk yield, etc etc) based on the year, and we keep ALL pedigree and ALL markers. A more refined approach is to keep pedigree and markers only up to the same date, for instance if "partial"= March 2014 we should keep records, pedigree and markers up to March 2014 (because pedigree and markers were used to predict the young rams).

6. In genetic evaluations with Unknown Parent Groups, the EBVs are not estimable functions So you need to refer all EBVs to a common genetic base in order to infer "bias" or not. Typically the genetic base is something like "average EBV of all females born in 2010" or something like that.

All this requires good knowledge of the data sets, the breeding scheme (or the breed), and a good command of scripting and genetic

4.4 The importance of several comparisons

The Figure 1 below shows all the estimates of b_{pw} in F. L. Macedo et al. (2020). For instance, in the X-axis we see the year of cut-off of *partial*, and the repeated points correspond to several *whole* years: 2010, 2011… It is clear that there is a large variation of b_{pw} due to chance, so to assess the unbiasedness of genetic evaluation one should do several pairs of *whole* and *partial* and not rely on a single study. For instance, year 2008 evaluation was clearly biased ($b_{pw} < 1$) whereas the other years were not.



Different estimates of b_{pw}

4.5 Estimation of genomic accuracies vs pedigree ones

How do I infer if a genomic evaluation is more accurate than a pedigree based one? There are two manners.

The first approach is to use *whole* and *partial* as we have explained so far, and evaluate each run both BLUP and genomic prediction (e.g. SSGBLUP), which yields a Table like above. This gives quite complete information as we can compare accuracies across methods and at different times.

The second approach is to consider that the genomic evaluation has "more data" so the pedigree-based evaluation is *partial* and the genomic evaluation is *whole*. The records $\mathbf{y}$ are the same. in both. Then the statistics above describe the ratio, increase, or absolute accuracies. For instance if we observe $\hat{\rho}_{pw} = 1$, it means that adding genotypes did not change anything. However, if we obtain $\hat{\rho}_{pw} = 0.9$, it means that accuracy increased (relatively) by $\frac{1}{\hat{\rho}_{pw}} - 1 = 0.11$.

5 Contribution of Smarter to this Deliverable and publications

During project Smarter, we verified the quality of the procedure using simulation in Macedo et al. (2020a), derived the estimator of "Unselected Accuracy" in Macedo et al. (2020b, 2021) and applied the method to two breeds of dairy sheep: Manech Tete Rousse and Lacaune, with extensive testing of different models for genetic evaluation and of different traits (Macedo et al., 2020b; Macedo et al., submitted).

Macedo, FL, A Reverter, and Andrés Legarra. 2020a. “Behavior of the Linear Regression Method to Estimate Bias and Accuracies with Correct and Incorrect Genetic Evaluation Models.” *Journal of Dairy Science* 103 (1): 529–44.

Macedo, Fernando L., Ole F. Christensen, Jean-Michel Astruc, Ignacio Aguilar, Yutaka Masuda, and Andrés Legarra. 2020b. “Bias and Accuracy of Dairy Sheep Evaluations Using BLUP and SSGBLUP with Metafounders and Unknown Parent Groups.” *Genetics, Selection, Evolution : GSE* 52 (August). https://doi.org/10.1186/s12711-020-00567-1.

Macedo, Fernando L., Ole F. Christensen, and Andrés Legarra. 2021. “Selection and Drift Reduce Genetic Variation for Milk Yield in Manech Tête Rousse Dairy Sheep.” *JDS Communications* 2 (1): 31–34. https://doi.org/10.3168/jdsc.2020-0010.

Submitted: F. L. Macedo, J.M. Astruc, T.H.E. Meuwissen and A. Legarra “Removing data and using metafounders alleviates biases for all traits in Lacaune dairy sheep predictions”

6 References

Bermann, Matias, Andres Legarra, Mary Kate Hollifield, Yutaka Masuda, Daniela Lourenco, and Ignacy Misztal. 2021. “Validation of Single-Step GBLUP Genomic Predictions from Threshold Models Using the Linear Regression Method: An Application in Chicken Mortality.” *Journal of Animal Breeding and Genetics* 138 (1): 4–13. https://doi.org/https://doi.org/10.1111/jbg.12507.

Bijma, P. 2012. “Accuracies of Estimated Breeding Values from Ordinary Genetic Evaluations Do Not Reflect the Correlation Between True and Estimated Breeding Values in Selected Populations.” *Journal of Animal Breeding and Genetics* 129 (5): 345–58.

Dekkers, J. C. M. 1992. “Asymptotic Response to Selection on Best Linear Unbiased Predictors of Breeding Values.” *Animal Science* 54 (3): 351–60. https://doi.org/10.1017/S0003356100020808.

Henderson, Charles R. 1975. “Best Linear Unbiased Estimation and Prediction Under a Selection Model.” *Biometrics*, 423–47. http://www.jstor.org/stable/2529430.

Henderson, C. R. 1982. “Best Linear Unbiased Prediction in Populations That Have Undergone Selection.” In *Proceedings of the World Congress on Sheep and Beef Cattle Breeding*, 1:191–201.

Legarra, Andres, and Antonio 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): 53.

Macedo, Fernando L., Ole F. Christensen, Jean-Michel Astruc, Ignacio Aguilar, Yutaka Masuda, and Andrés Legarra. 2020. “Bias and Accuracy of Dairy Sheep Evaluations Using BLUP and SSGBLUP with Metafounders and Unknown Parent Groups.” *Genetics, Selection, Evolution : GSE* 52 (August). https://doi.org/10.1186/s12711-020-00567-1.

Macedo, Fernando L., Ole F. Christensen, and Andrés Legarra. 2021. “Selection and Drift Reduce Genetic Variation for Milk Yield in Manech Tête Rousse Dairy Sheep.” *JDS Communications* 2 (1): 31–34. https://doi.org/10.3168/jdsc.2020-0010.

Macedo, FL, A Reverter, and Andrés Legarra. 2020. “Behavior of the Linear Regression Method to Estimate Bias and Accuracies with Correct and Incorrect Genetic Evaluation Models.” *Journal of Dairy Science* 103 (1): 529–44.

Reverter, A., B. L. Golden, R. M. Bourdon, and J. S. Brinks. 1994. “Technical Note: Detection of Bias in Genetic Predictions.” *Journal of Animal Science* 72 (1): 34–37. https://doi.org/10.2527/1994.72134x.

Sorensen, Daniel, R. Fernando, and D. Gianola. 2001. “Inferring the Trajectory of Genetic Variance in the Course of Artificial Selection.” *Genetical Research* 77 (01): 83–94. http://journals.cambridge.org/abstract_S0016672300004845.

7 Deviations or delays

No deviations

8 Acknowledgements

Fernando Macedo (Universidad de la Republica / INRAE / SLU) and Toni Reverter (CSIRO) contributed to most of the results in this document.