Session II

Validation

ITB Working group for validation tests of genetic and genomic evaluations

Working Group was assigned 2018

Members

Esa Mäntysaari

Zengting Liu

Peter Sullivan

Raphael Mrode

Paul VanRaden

Valentina Palucci

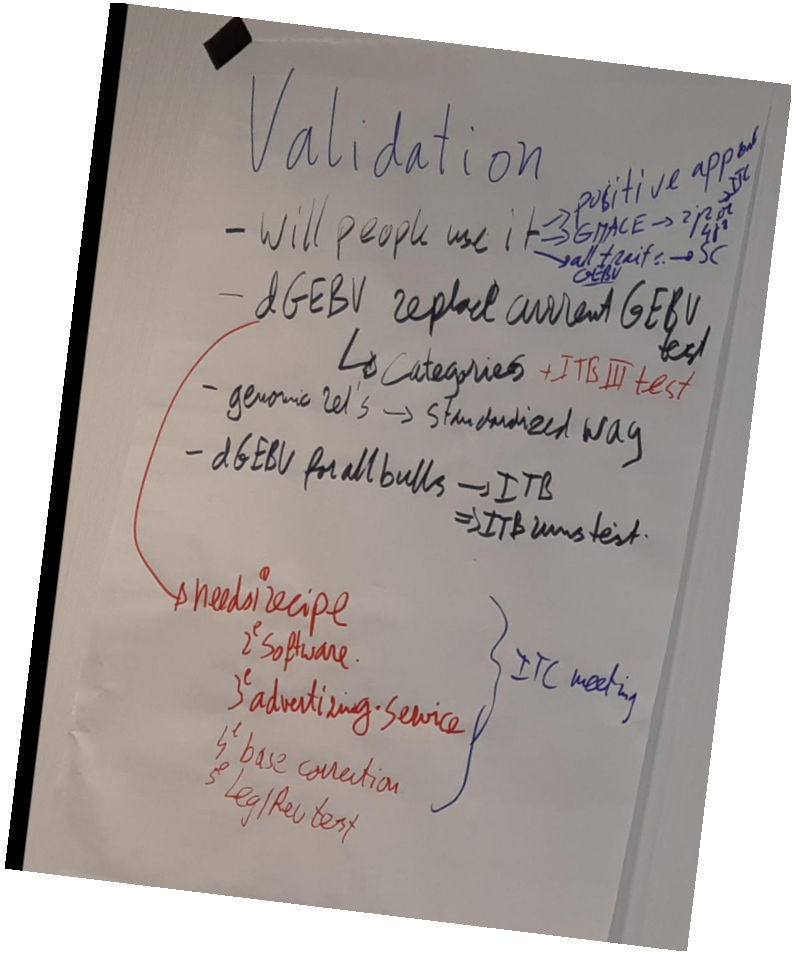
ITC Strategic Workshop 26.- 27.5.2022 - Summary of validation test discussions

The need for EBV/GEBV validation tests is still important:

1) Inclusion of national genetic evaluations to MACE
- the trend tests II and III and the MS trend test

2) Inclusion of genomic evaluations to GMACE (GEBVtest)
- Unbiasedness, inflation

3) Certification of genomic evaluations (GEBVtest)
- Unbiasedness, inflation, and accuracy
- To be used in international trade; required also by EU



Trend validation tests II, III, and the MS variance test

Applicability of the tests has not changed in genetic evaluations of countries that are NOT using genomics in selection

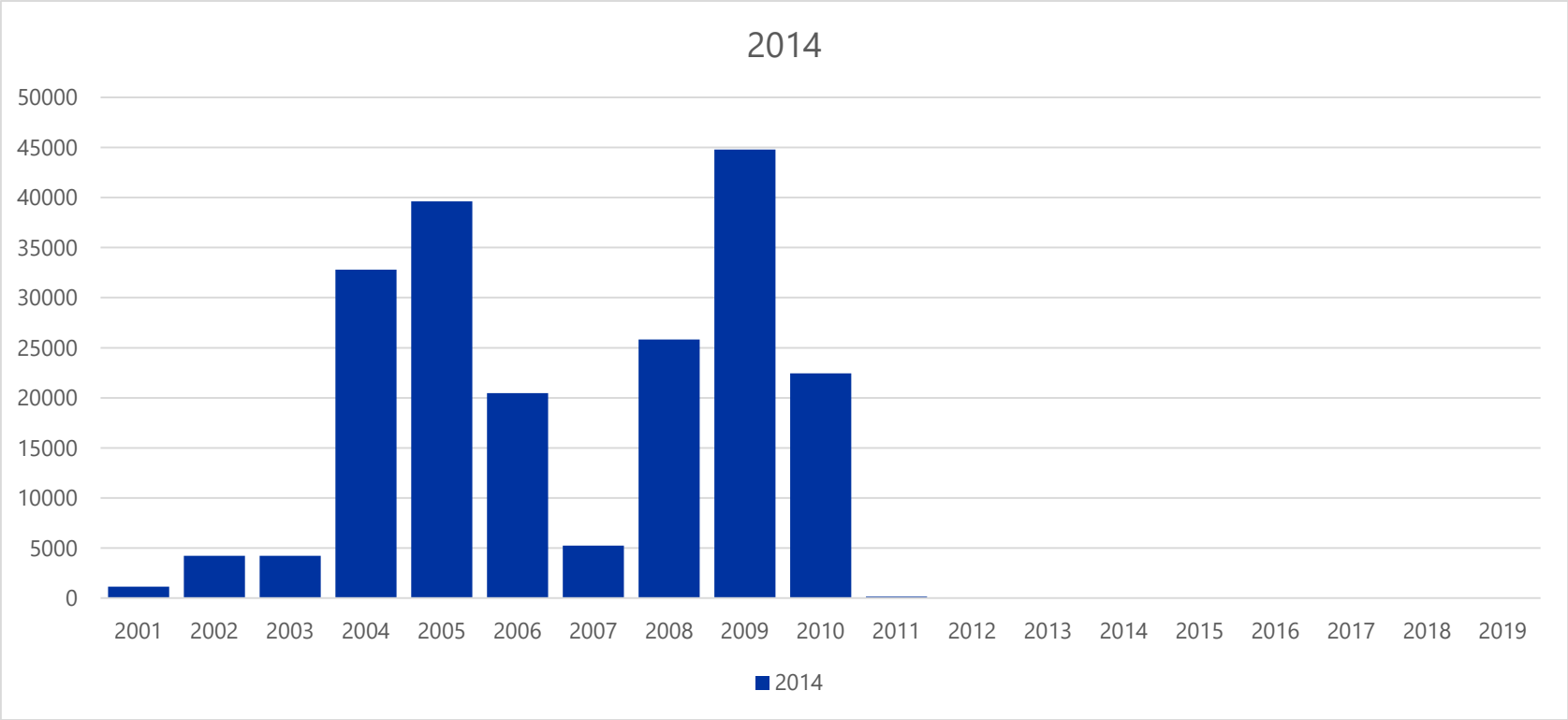
In countries applying genomic selection:

1) **tests for EBV**:

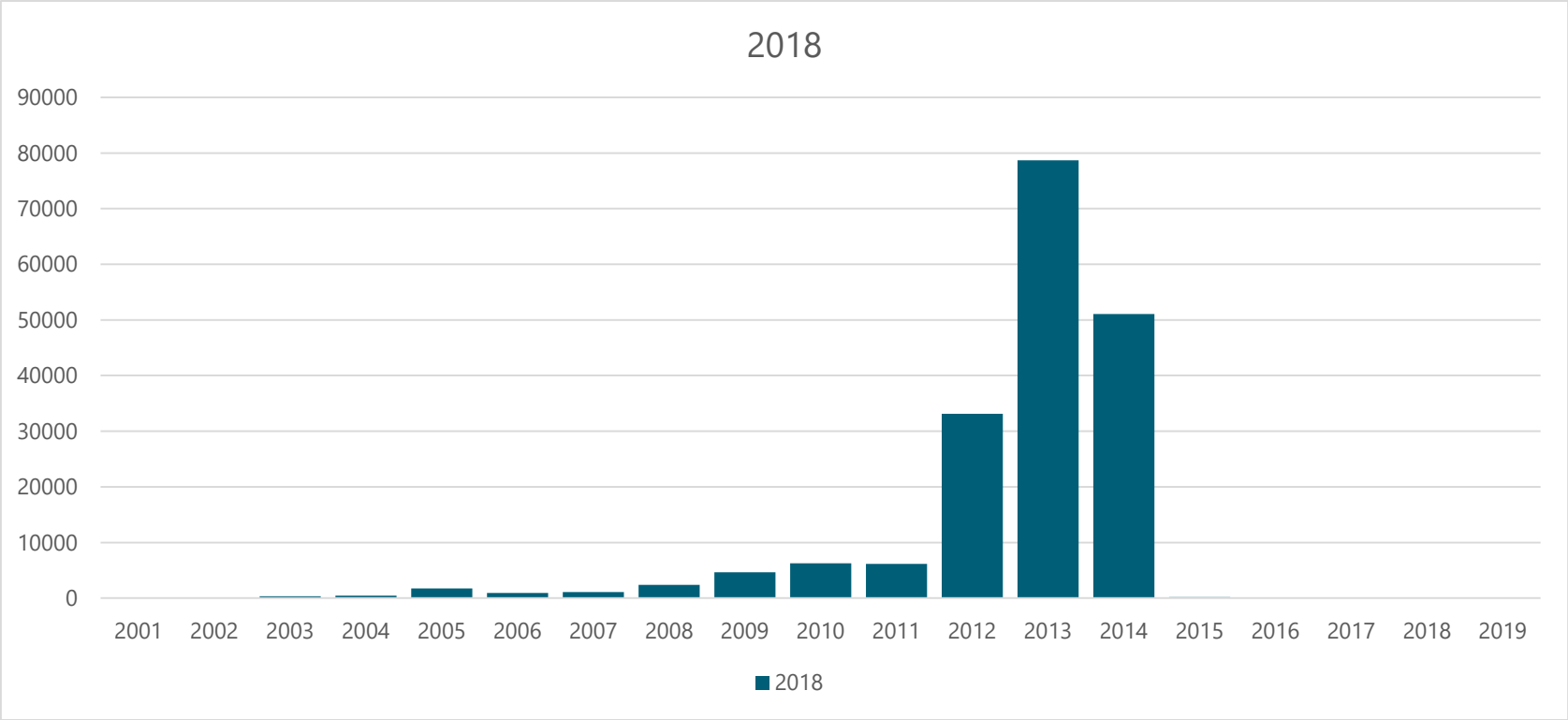
- All the trend tests apply as before, except that the power has been reduced because the sires are used only a short time
- Especially Test III has lost its power
- Therefore the test II, based on DYD, should be preferred, but the DYD for the test should be approximated rationally

In the case of efficient GS
- the EBV are likely to fail the tests!

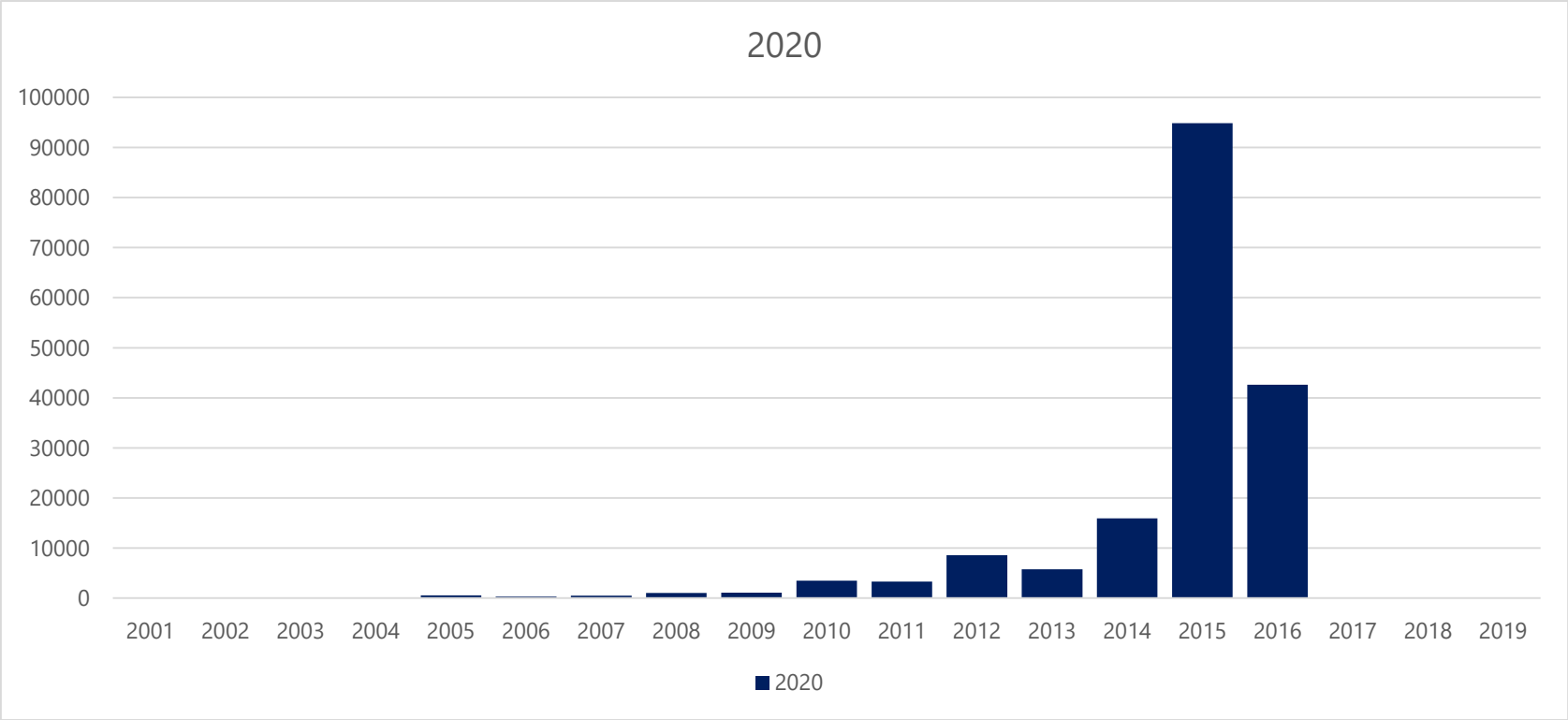
Distribution of sire birth years in DFS Holsteins calving first time year 2014



Distribution of sire birth years in DFS Holsteins calving first time year 2018



Distribution of sire birth years in DFS Holsteins calving first time year 2020



Number of test submitted to Interbull during last 3 years (TRAIT MILK)

Tests applied	2020	2021	2022
Test II	4	1	1
Test III	8	6	6
GEBVtest	2	3	2

[§]Valentina Palucci, Interbull 2023

Very difficult to conclude anything about the changes in results

2) **tests to be applied on GEBV**

Countries own interest:

- Validation test II can be used for single-step evaluations
- Validation test III can be modified to use GEBVs and R^2 of GEBV

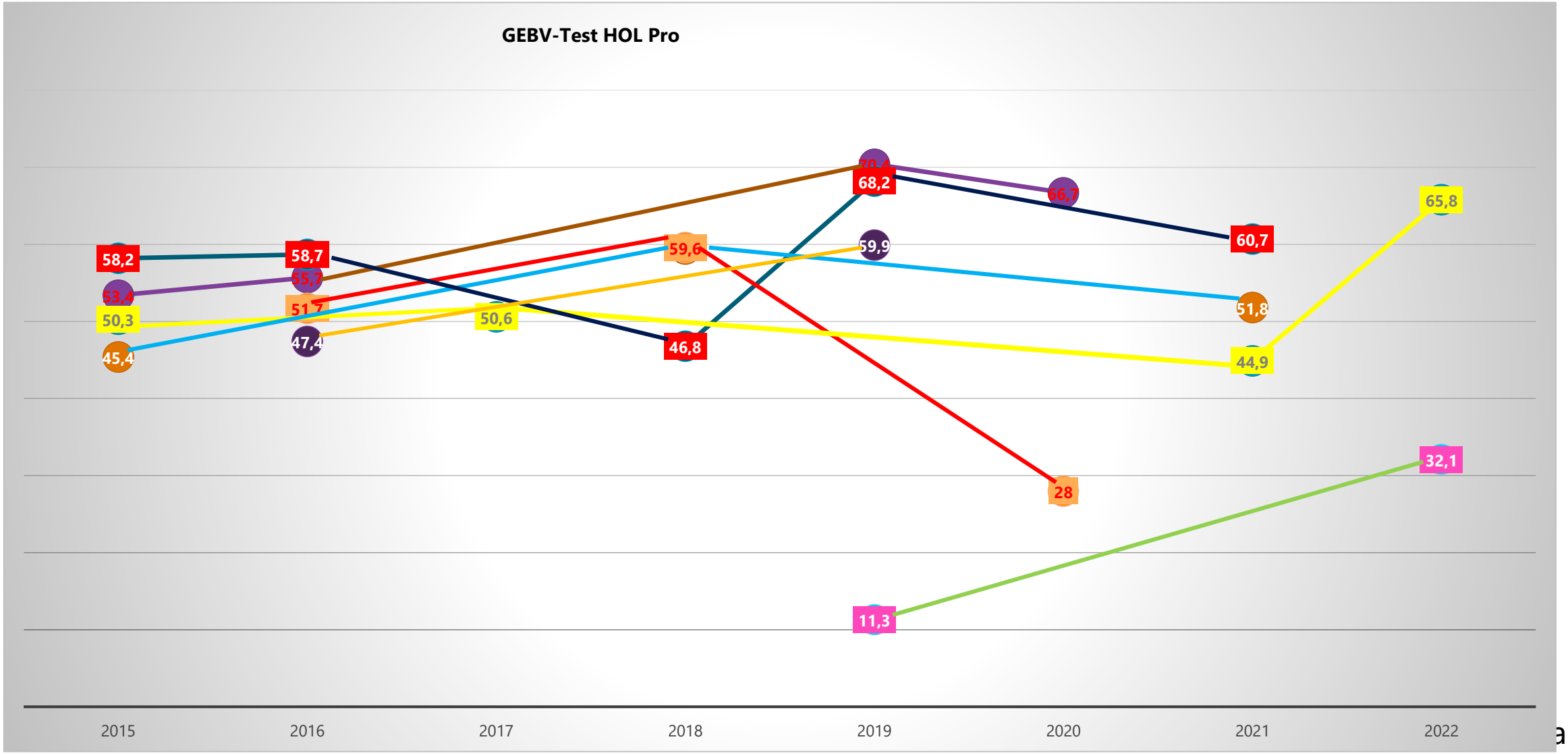
International evaluations (for exporting semen, or for GMACE input)

- Interbull GEBV test
  - Tests how well the early GEBV predicts (recent) progeny based EBVs
  - To pass: regression coefficient b1 ~ 1.0 and $R^2_{GEBV} > R^2_{PA}$

Number of test submitted to Interbull during last 3 years (TRAIT MILK)

Tests applied	2020	2021	2022
Test II	4	1	1
Test III	8	6	6
GEBVtest	**2**	**3**	**2**

GEBVtest example:
R^2 results of countries having applied test >2 times (HOL protein)

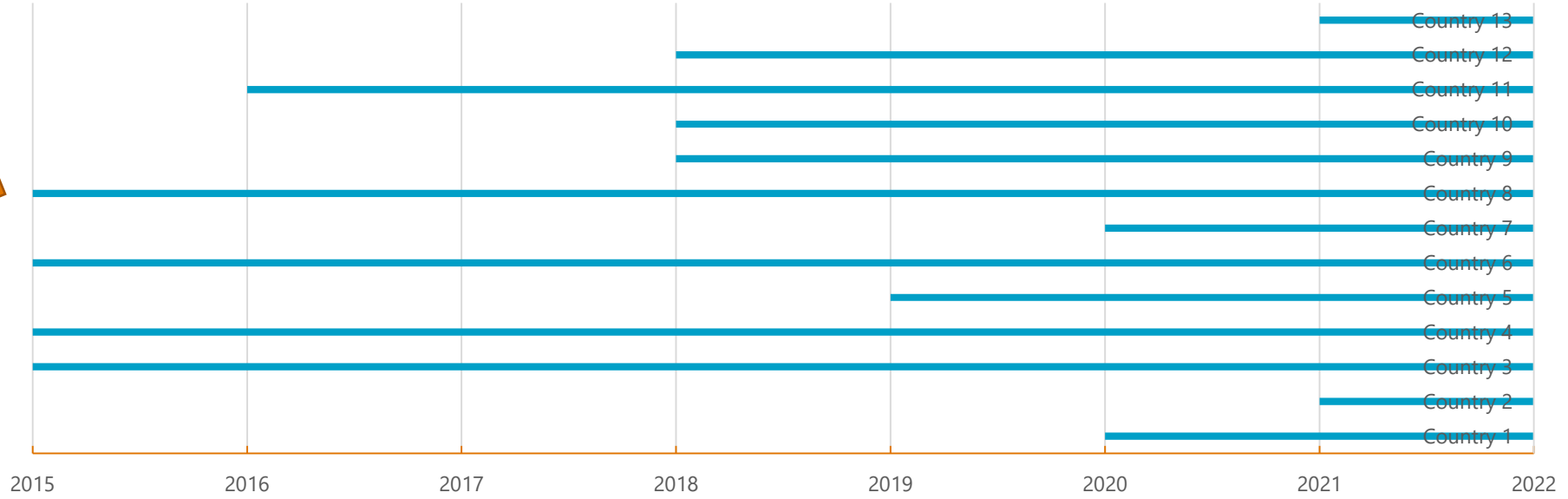


Number of test submitted to Interbull during the last 3 years (TRAIT MILK)

Tests applied	2020	2021	2022
Test II	4	1	1
Test III	8	6	6
GEBVtest	**2**	**3**	**2**

Age of the last accepted GEBVtest in 13 countries

4 countries have 8 years old results



Weak points in the current GEBVtest

Can we assume that the

DYD or DRP from the EBV evaluations are unbiased?

- Genomic evaluations appear biased
  if the recent EBV of the validation bulls is lower than the GEBV
- Genomic evaluations appear inflated
  if the recent EBV has less variance than the variance in GEBV
- Genomic evaluation accuracy (R^2) is under-estimated
  if EBV are biased

ITC Strategic Workshop 26.- 27.5.2022 - validation test conclusions

2) **tests to be applied on GEBV** :

GEBVtest should not assume EBV evaluations to be a "golden standard"

Better approaches:

- Testing using two consecutive GEBVs : $GEBV_{red}$ and $GEBV_{full}$
- Test of changes in GEBV i.e. deregressed GEBV (dGEBV)

To standardize the testing, these new options were included into current GEBVtest program → GEBVtest2022.py

Thank you!

Ja kiitos hyvistä vuosista!

Acknowledgements:

Valentina Palucci compiled the statistics of current tests

Validation working group

Peter Sullivan for the new version of GEBVtest2022
NAV provided HOL calving data

theory & actual experiences with method LR for validation tests

Andrés Legarra, andres.legarra@uscdcb.com

Interbull Technical Workshop, Rome, 14-15 Feb 2023

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

- Intro & why
- how does it work (theory & practicalities)
- examples of output

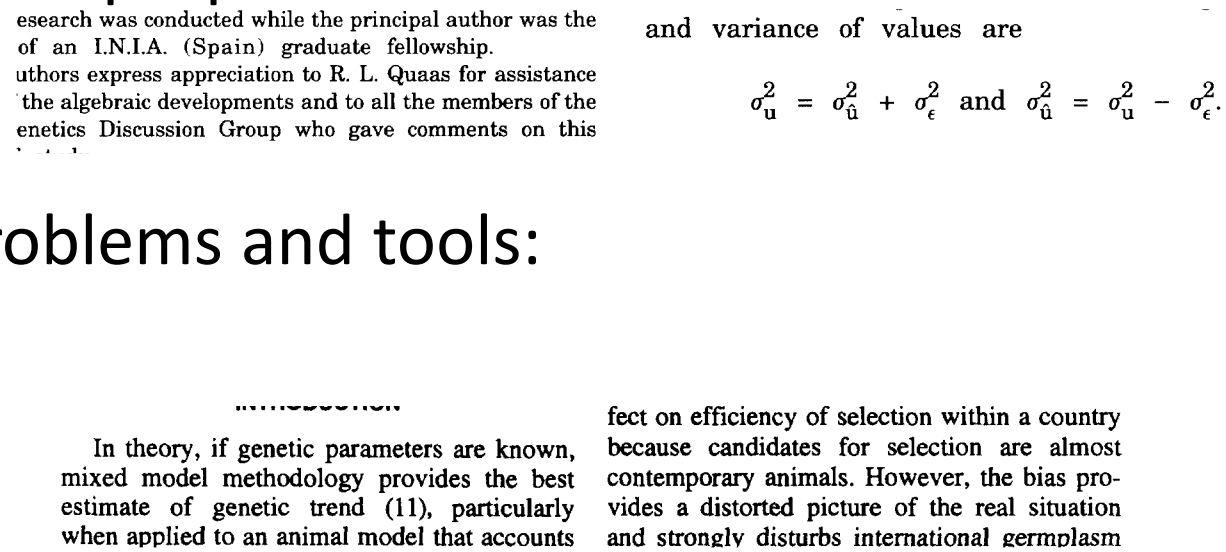
Motivations

- Esa Mantysaari invited me to give this talk (before ☺ he knew that I now work for CDCB)
- I show theory and experiences that we got from the use of the LR method (mainly in dairy sheep from France)
- There is a tutorial: http://genoweb.toulouse.inra.fr/~alegarra/SMARTER_D5.3_Use_of_method_LR.pdf
- and in the course notes: http://genoweb.toulouse.inra.fr/~alegarra/GSIP.pdf

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Comparing models, history

- In 1994 Toni Reverter (then in Colorado) and Didier Boichard (INRA) attacked the problem of checking bias in BLUP evaluations
- Reverter focused on “Hendersonian » properties of BLUP:
  - comparing succesive evaluations
  - more theorems
- Boichard focused on dairy cattle problems and tools:
  - comparing first-lactation and all-lactation
  - check of DYDs
  - comparing succesive evaluations



esearch was conducted while the principal author was the
of an I.N.I.A. (Spain) graduate fellowship.
uthors express appreciation to R. L. Quaas for assistance
the algebraic developments and to all the members of the
enetics Discussion Group who gave comments on this

and variance of values are

$$\sigma^2_u = \sigma^2_{\hat{u}} + \sigma^2_\epsilon \text{ and } \sigma^2_{\hat{u}} = \sigma^2_u - \sigma^2_\epsilon.$$

In theory, if genetic parameters are known, mixed model methodology provides the best estimate of genetic trend (11), particularly when applied to an animal model that accounts

fect on efficiency of selection within a country because candidates for selection are almost contemporary animals. However, the bias provides a distorted picture of the real situation and strongly disturbs international germplasm

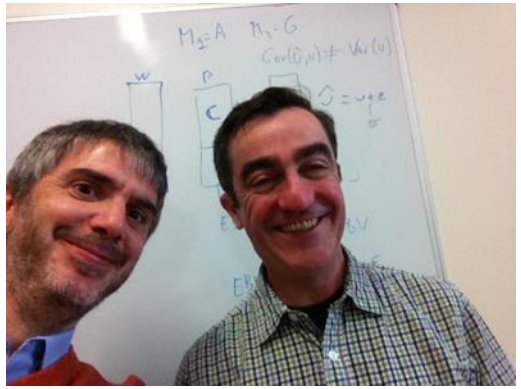
CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Why method LR

- Later we had the explosion of cross-validation due to Genomics
- My own work in species with small offspring size (pigs, sheep)
- In these cases DYDs do not exist or… getting DYDs right is very hard !!
  - are dams well estimated? are contemporary groups well estimated? …etc
- Should we trust DRP?
  - I don’t trust Reliabilities (approximations)
  - Equations for DRPs are often also approximations (even if Reliabilities are correct)
  - I trust (G)EBVs

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Why method LR



- We wanted something more general
- Toni and I made mutual visits
- We re-derived his 1994 paper to apply to “multiple individual” case (the 1994 one considered individual EBVs)
- In this manner we derived more general equations that explicitly show relationships, Prediction Error Covariances, and also effects of selection
- We also focused on WHICH properties we want to check and WHY

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Metrics

- Theory of quantitative genetics suggest using Metrics from linear regression of u (TBV) on $\hat{u}$ (EBV) for a *collection* of animals (and hence the vectors)

- Bias: $\Delta = \frac{1}{n}(\mathbf{1}'\hat{\mathbf{u}} - \mathbf{1}'\mathbf{u})$ (it is NOT the intercept of the regression of u on $\hat{u}$)
- Slope: $b = \frac{Cov(\mathbf{u},\hat{\mathbf{u}})}{Var(\hat{\mathbf{u}})}$ (slope of the regression of u on $\hat{u}$)
- Accuracy: $r = \frac{Cov(\mathbf{u},\hat{\mathbf{u}})}{\sqrt{Var(\mathbf{u})Var(\hat{\mathbf{u}})}}$

In fact: $MSE = \frac{1}{n}\sum_i (u_i - \hat{u}_i)^2 = \Delta^2 + Var(\mathbf{u})\left(1 + \frac{r^2}{b^2} - \frac{2r^2}{b}\right)$

- Why are these relevant? Genetic progress !!

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

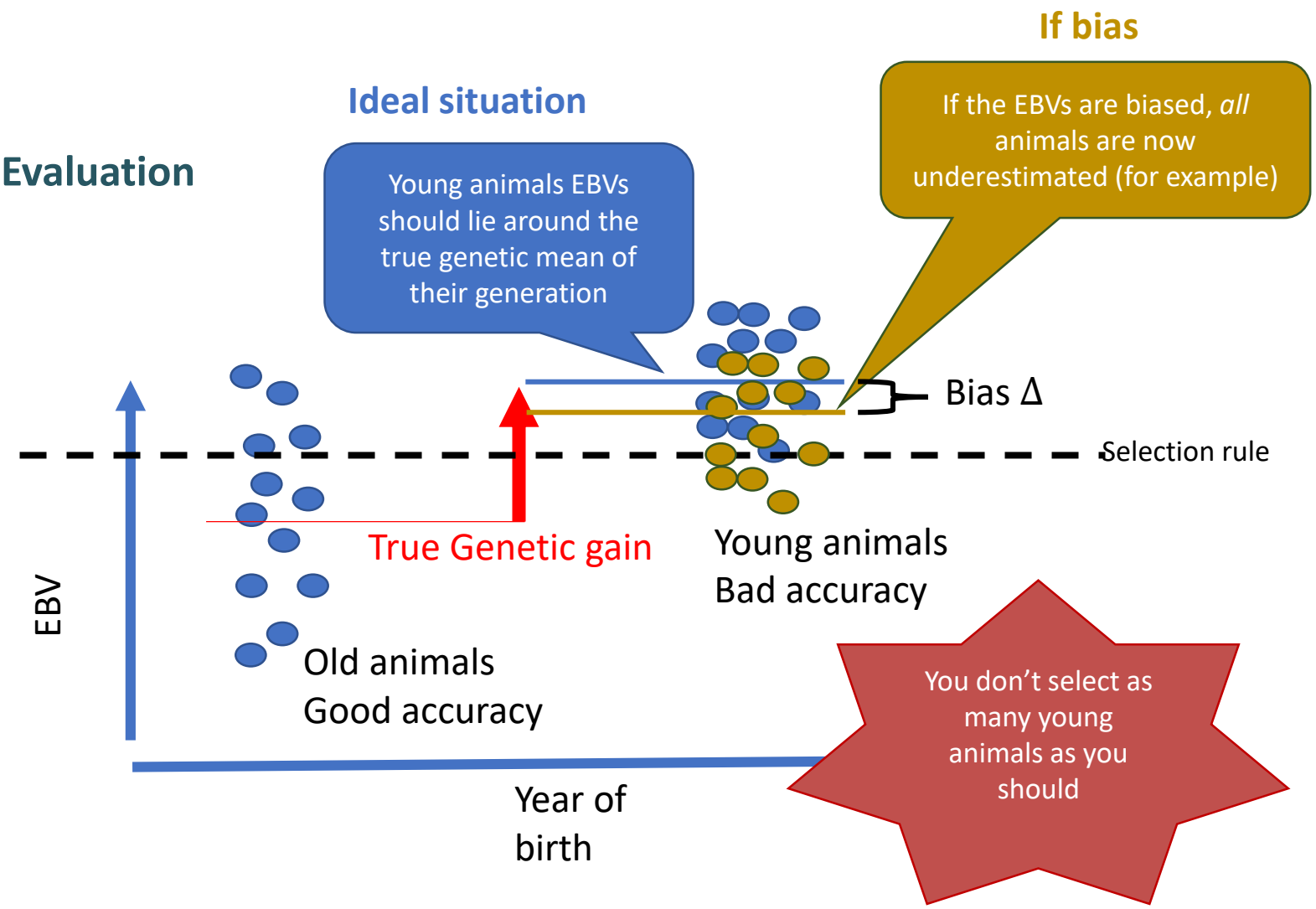
True and estimated genetic progress

- When we select animals, we *believe* our $\Delta G = \frac{1}{n}\Sigma(GEBV) = \bar{\hat{u}}$ of selected animals
- This only holds if bias $\Delta = 0$, regression $b = 1$
- $\Delta > 0$ (bias) or $b < 1$ (overdispersion) lead to too high values of selected young animals
- Both need to be checked
  - $\Delta > 0$ implies that all animals were overevaluated (regardless of $b = 1$)
  - $b < 1$ implies that selected animals are overevaluated (regardless of $\Delta = 0$)

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

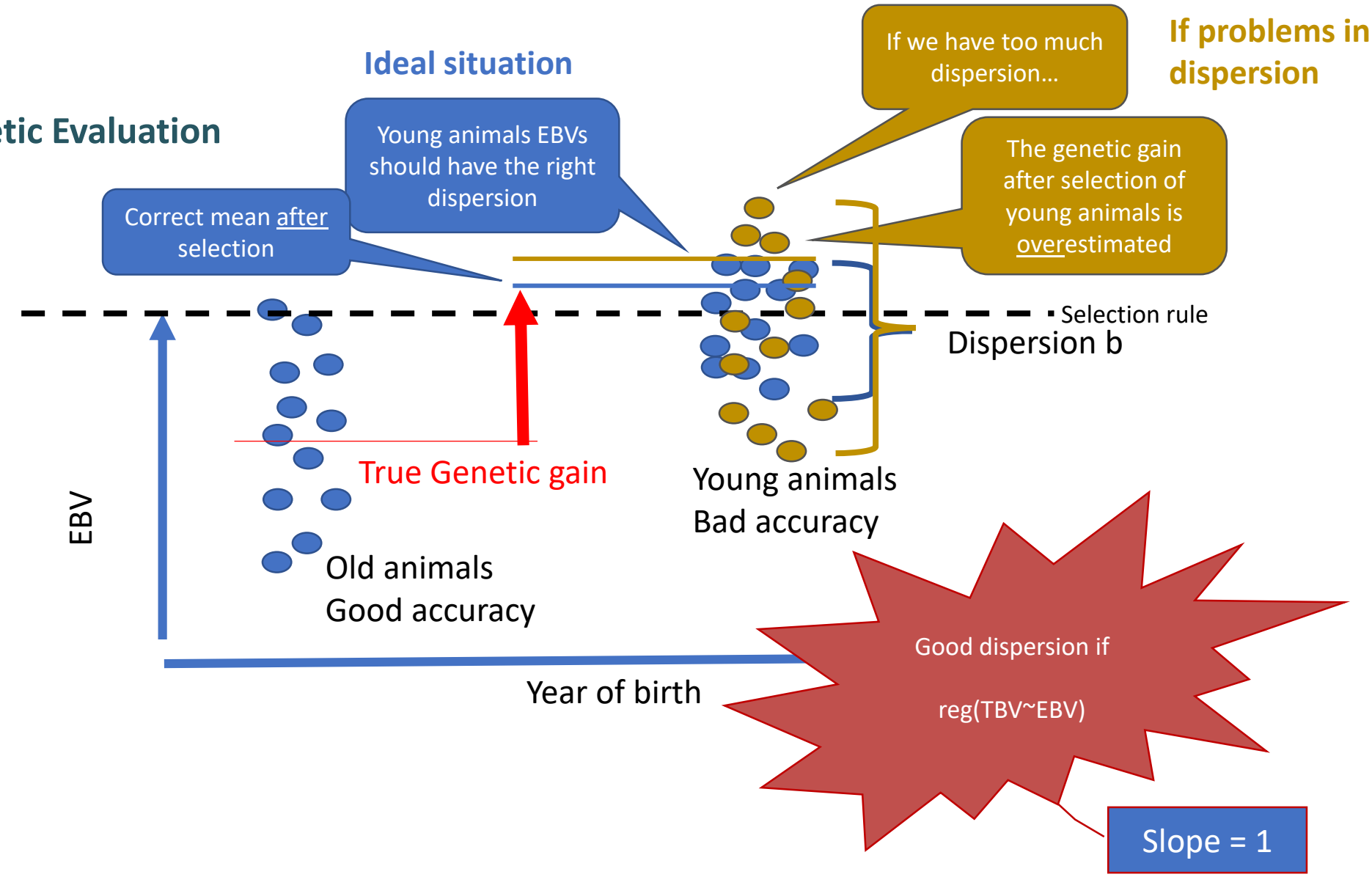
Genetic gain: Δ

Consider a Genetic Evaluation

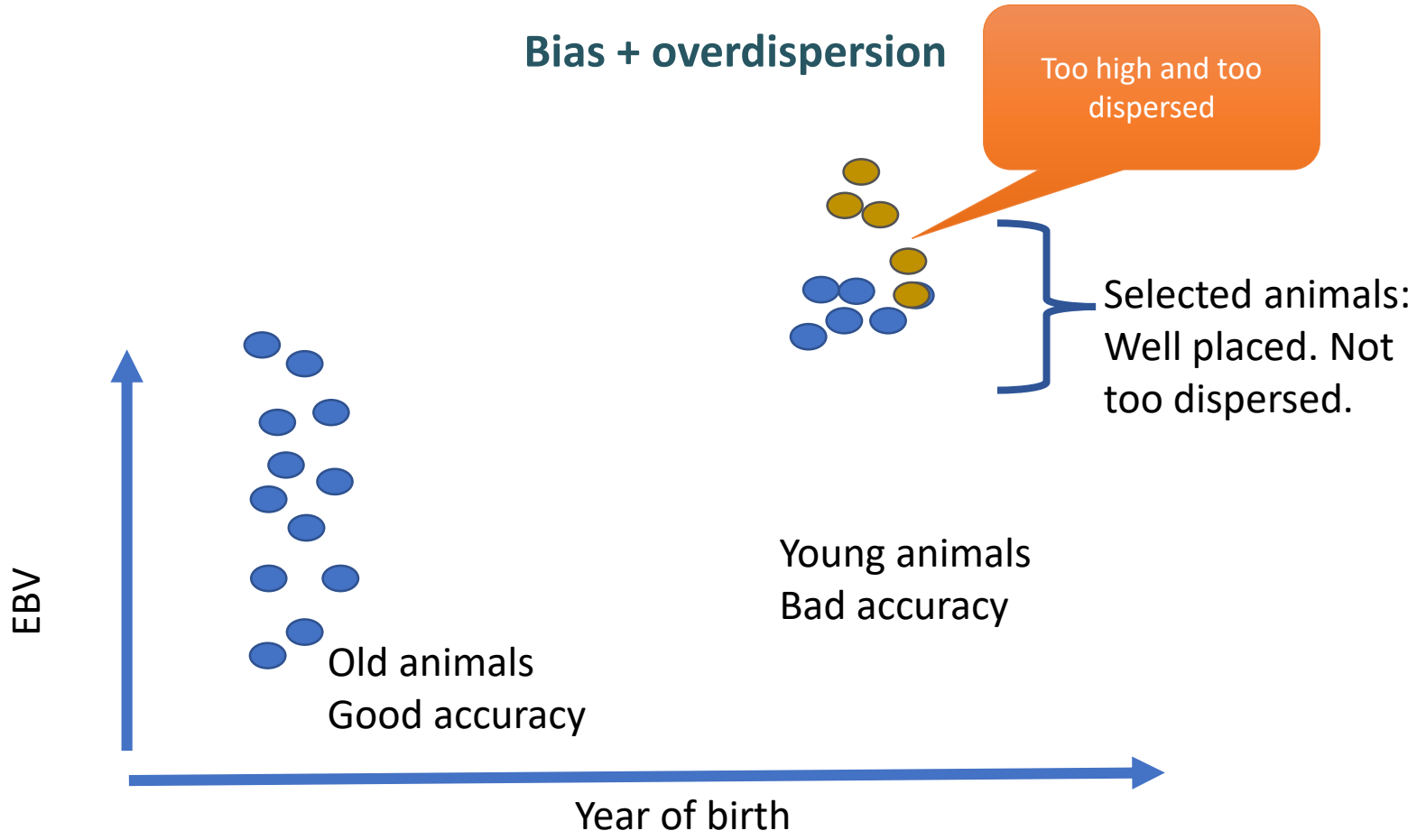


Dispersion: b

Consider a Genetic Evaluation



Bias + overdispersion



We don’t have true EBVs

yet Legarra & Reverter (2018) proposed a new method based on comparisons of EBV from “partial” (old) data vs “whole” (old+new) data.

- Does not require “true” breeding values
- Does not require pre-corrected phenotypes
- Could be used for any kind of traits

- Legarra, A., & Reverter, A. (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), 1-18.
- Legarra, A., & Reverter, A. (2019). Correction to: Semi-parametric estimates of population accuracy and bias of predictions of breeding values and future phenotypes using the LR method. *Genetics Selection Evolution*, *51*(1), 1-2.

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

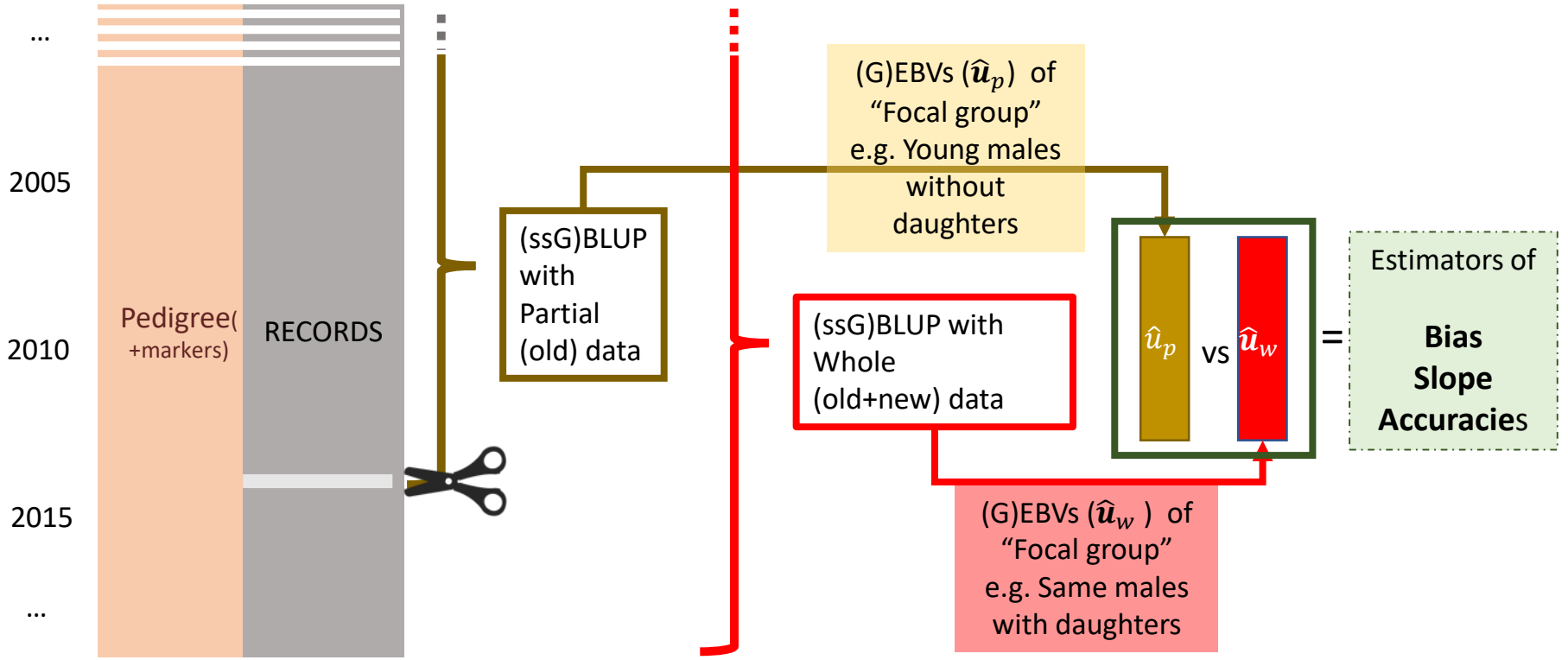
Check of bias using successive evaluations

- We proved (analytically) that in successive genetic evaluations there are useful statistical properties of the JOINT distributions of “early”(partial) and “late” (whole) EBVs
- We use these properties to get estimators of biases and accuracies
- Some details in the paper referring to accuracies (not to biases) were later refined
  - series of papers by Macedo et al. and the Tutorial
  - don’t worry about them today

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

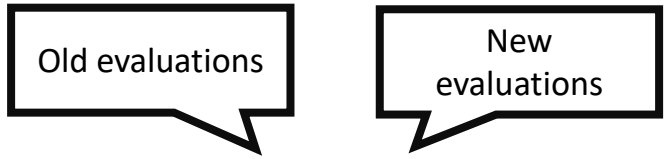
The proposed method LR

How does LR method work?



CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Estimators of LR method: Bias and Slope



Bias $\hat{\Delta}_p = \bar{\hat{\boldsymbol{u}}}_p - \bar{\hat{\boldsymbol{u}}}_w$

Expected value of 0 in absence of bias

Slope $\hat{b}_{w,p} = \frac{cov(\hat{\mathbf{u}}_p, \hat{\mathbf{u}}_w)}{var(\hat{\mathbf{u}}_p)}$

Expected value of 1 in unbiased genetic evaluations

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

2018

Top 50 Ranked by Net Merit $

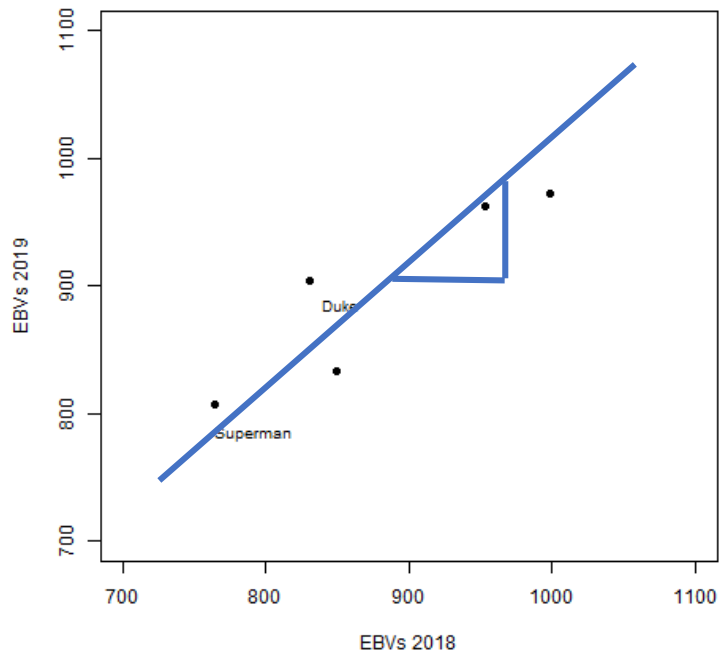
NAAB	Name	NM$	Rel.
29HO17553	JOSUPER	999 G	96
1HO11881	PRINCETON	849 G	93
250HO13267	DUKE	831 G	82
151HO681	RUBICON	953 G	95
200HO7846	SUPERMAN	764 G	93

2019

Top 50 Net Merit $

NAAB Code	Short Name	NM$	OR	Rel
29HO17553	JOSUPER *99-I	973	G	97
1HO11881	PRINCETON *99-I	833	G	95
250HO13267	DUKE	904	G	92
151HO681	RUBICON	963	G	97
200HO7846	SUPERMAN	807	G	95

$$\hat{\Delta}_p = \frac{999 + 849 + 831 + 953 + 764}{5} - \frac{973 + 833 + 904 + 963 + 807}{5} = -16.8$$




```
EBV2018=c(999,849,831,953,764)
EBV2019=c(973,833,904,963,807)
delta_p=mean(EBV2018)-mean(EBV2019) # -16.8   Δ̂p
aa=lm(EBV2019~EBV2018)
b_wp=aa$coefficients[2] # 0.71   b̂w,p
```

Some of the algebra

$$Var\begin{pmatrix}\hat{\mathbf{u}}_p\\ \hat{\mathbf{u}}_w\\ \mathbf{u}\end{pmatrix} = \begin{pmatrix}\mathbf{G}-\mathbf{C}_p^{uu} & \mathbf{G}-\mathbf{C}_p^{uu} & \mathbf{G}-\mathbf{C}_p^{uu}\\ \mathbf{G}-\mathbf{C}_p^{uu} & \mathbf{G}-\mathbf{C}_w^{uu} & \mathbf{G}-\mathbf{C}_w^{uu}\\ \mathbf{G}-\mathbf{C}_p^{uu} & \mathbf{G}-\mathbf{C}_w^{uu} & \mathbf{G}\end{pmatrix}$$

$$E\left(\frac{1}{n}\left(\hat{\mathbf{u}}_p-\bar{\hat{\mathbf{u}}}_p\right)'\left(\hat{\mathbf{u}}_w-\bar{\hat{\mathbf{u}}}_w\right)\right)$$
$$= \frac{1}{n}tr\left(\mathbf{SG}-\mathbf{SC}_p^{uu}\right)$$
$$= \overline{diag(\mathbf{G})}-\bar{\mathbf{G}}-\left(\overline{diag\left(\mathbf{C}_p^{uu}\right)}-\overline{\left(\mathbf{C}_p^{uu}\right)}\right)$$

Regression of whole on partial

after algebra on expectation of quadratic forms, $E\left(\hat{b}_{w,p}\right)=1$

$$\hat{b}_{w,p}=\frac{cov\left(\hat{\mathbf{u}}_p,\hat{\mathbf{u}}_w\right)}{var\left(\hat{\mathbf{u}}_p\right)}=\frac{\frac{1}{n}\left(\hat{\mathbf{u}}_p-\bar{\hat{\mathbf{u}}}_p\right)'\left(\hat{\mathbf{u}}_w-\bar{\hat{\mathbf{u}}}_w\right)}{\frac{1}{n}\left(\hat{\mathbf{u}}_p-\bar{\hat{\mathbf{u}}}_p\right)'\left(\hat{\mathbf{u}}_p-\bar{\hat{\mathbf{u}}}_p\right)}.$$

- expressing everything as quadratic forms allows to derive properties.
- e.g. the s.e. of $\hat{b}_{w,p}$ (slope of "whole" on "partial"), will be a function of
  - (1) number of individuals (2) their relationships (3) their PEV and PEC.

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Practicalities: defining focal groups

- The properties of the method hold and are useful for a group of animals that are contemporaries and have
  - similar selection pressure
  - similar information at ”partial” (e.g. only Parent Average or PA+genomics)
  - and similar information at “whole” predictions (e.g. Parent Average +genomics + phenotype, or Parent Average + genomics + offspring, or…)
- we call this “focal group”
- e.g.
  - young genomic bulls vs. same bulls with daughters
  - 1st-calving cows vs. same cows at 2nd-calving

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Practicalities: defining “whole” and “partial”

- You can do many “partials” and many “wholes”
- 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 !!

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

for instance: work in MTR

Evaluations with data until 2005, until 2006 and so on until 2017.

We compare

- **EBVs at birth** (EBV_P) of a set of Artificial Insemination males (2005 – 2014)
- **EBVs of the same males in later evaluations** (after having progeny) (EBV_W) (until 2017).

For example for males born in 2005, 11 pairs of evaluations were analysed,

- 2005 vs 2007; 2005 vs 2008; ... and 2005 vs 2017

The same for males born in 2006, 2007 ... 2015

- 2006 vs 2008; ... and 2015 vs 2017

Total of 65 comparisons that we "averaged" using a linear model to account for unbalance (details in the Macedo et al. paper)

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Practicalities: referring to same genetic base

- in addition to the slope $\hat{b}_{w,p}$ of the regression $\hat{\boldsymbol{u}}_w \sim \hat{\boldsymbol{u}}_p$
- we <u>do need</u> to check bias in the genetic trends $\hat{\Delta}_p = \bar{\hat{\boldsymbol{u}}}_p - \bar{\hat{\boldsymbol{u}}}_w$
- In genetic evaluations with Unknown Parent Groups, the EBVs are not estimable functions
- So you need to refer EBVs in “whole” and in “partial” to the same genetic base in order to infer “bias”
- Typically the genetic base is something like “average EBV of all females born in 2010” or something like that.

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Some results

- Dairy sheep improvement is a French specialty !
- This is not meat or wool sheep
- Very well structured co-operative “mini – dairy cattle style” breeding program
- AI, performance recording, etc etc
  - <2015: progeny-testing (first crop ~30 daughters)
  - >2015: genomic selection
- All results here concern Milk Yield

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

The breeds

Manech Tête Rousse

Female population size:	274,000
Females in the breeding flocks:	80,260 (29%)
Tested rams per year:	150
Rams at AI Center:	600
Individuals in the pedigree	540,999
Number of records Milk Yiedl:	1,842,295
Missing pedigree:	≈25%

Traits selected: Milk Yield and contents

Lacaune

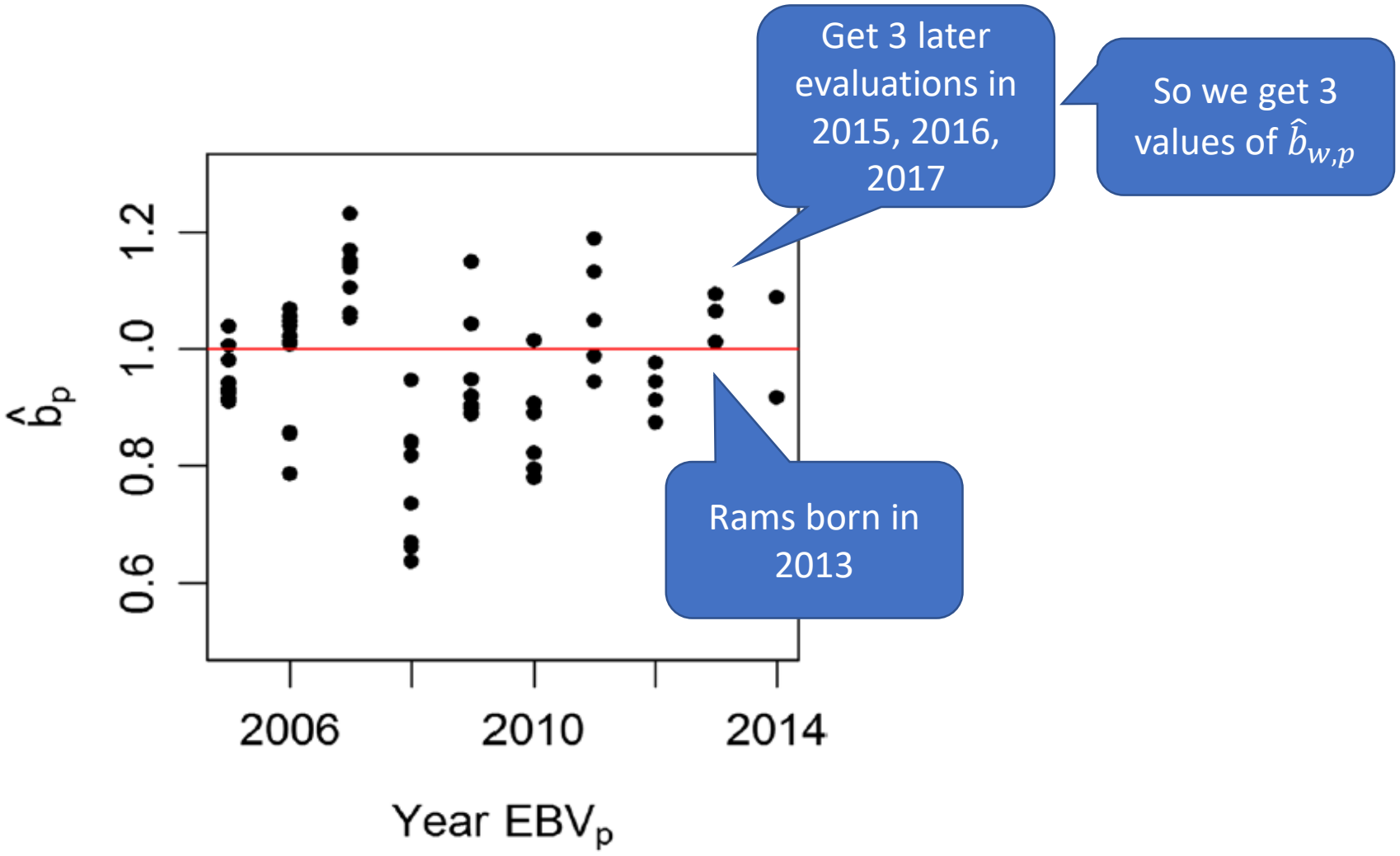
Female population size:	890,000
Females in the breeding flocks:	174,472 (19%)
Tested rams per year:	440
Rams at AI Center:	1400
Individuals in the pedigree	1,868,975
Number of records Milk Yield:	5,696,348
Missing pedigree:	≈9%

Traits: Milk Yield and contents, SCS, Udder traits

Main results Manech Tête Rousse

each focal group has ~150 rams

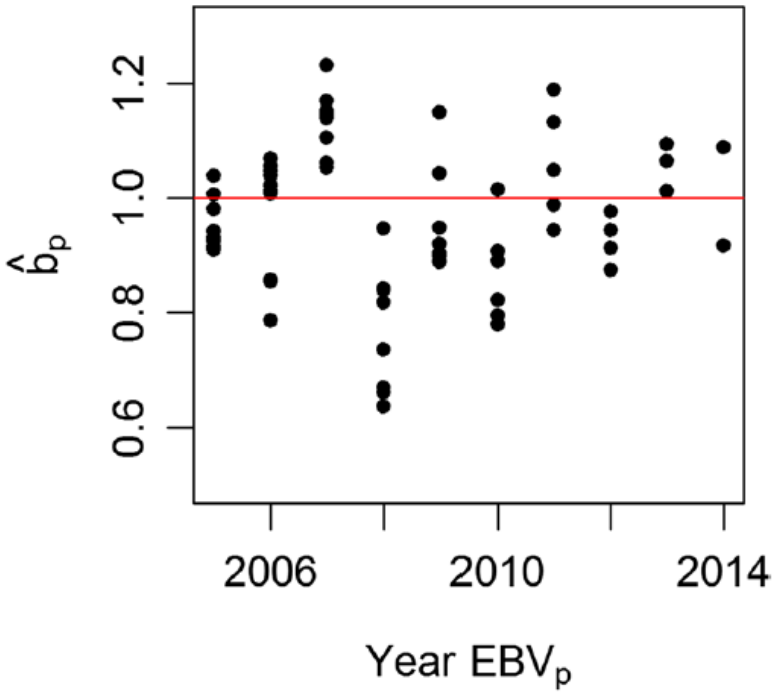
BLUP-MF



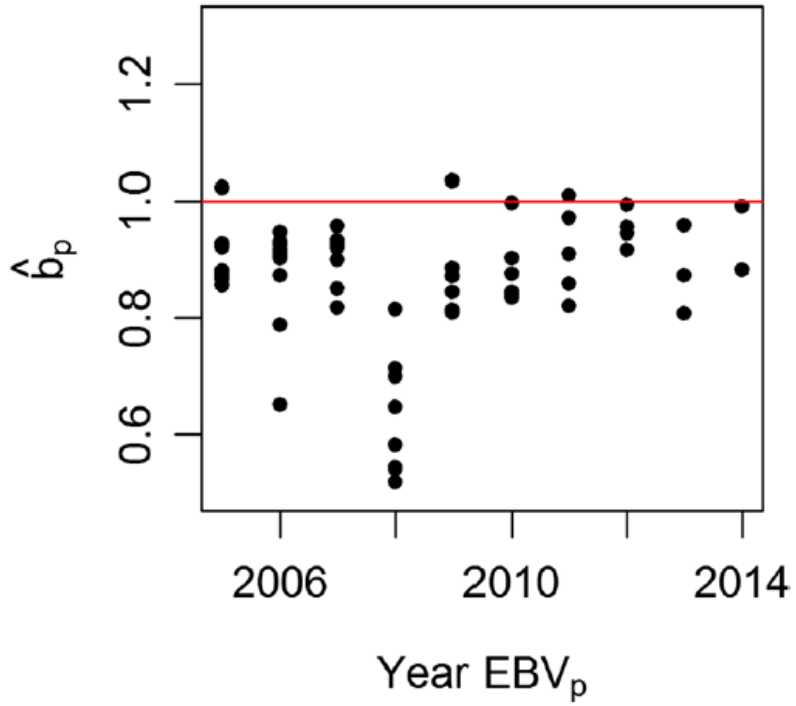
Main results Manech Tête Rousse

each focal group has ~150 rams

BLUP-MF



SSGBLUP-UPGH



Important variation between truncation points!

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Main results Manech Tête Rousse

Some models (UPGH…) are really biased

Very small bias in Manech Tête Rousse!

- slope $\hat{b}_{w,p}$~1
- small, positive bias $\hat{\Delta}_{p}$ (0.2 genetic s.d.)
- lambs are over-predicted

Model	$\hat{\Delta}_{p}$	$\hat{b}_{w,p}$
BLUP-MF	0.25	0.98
BLUP-UPGA	0.48	0.96
SSGBLUP-MF	0.23	0.97
SSGBLUP-UPGA	0.32	0.94
SSGBLUP-UPGH	0.48	0.88

Across models

slope $\hat{b}_{w,p}$:

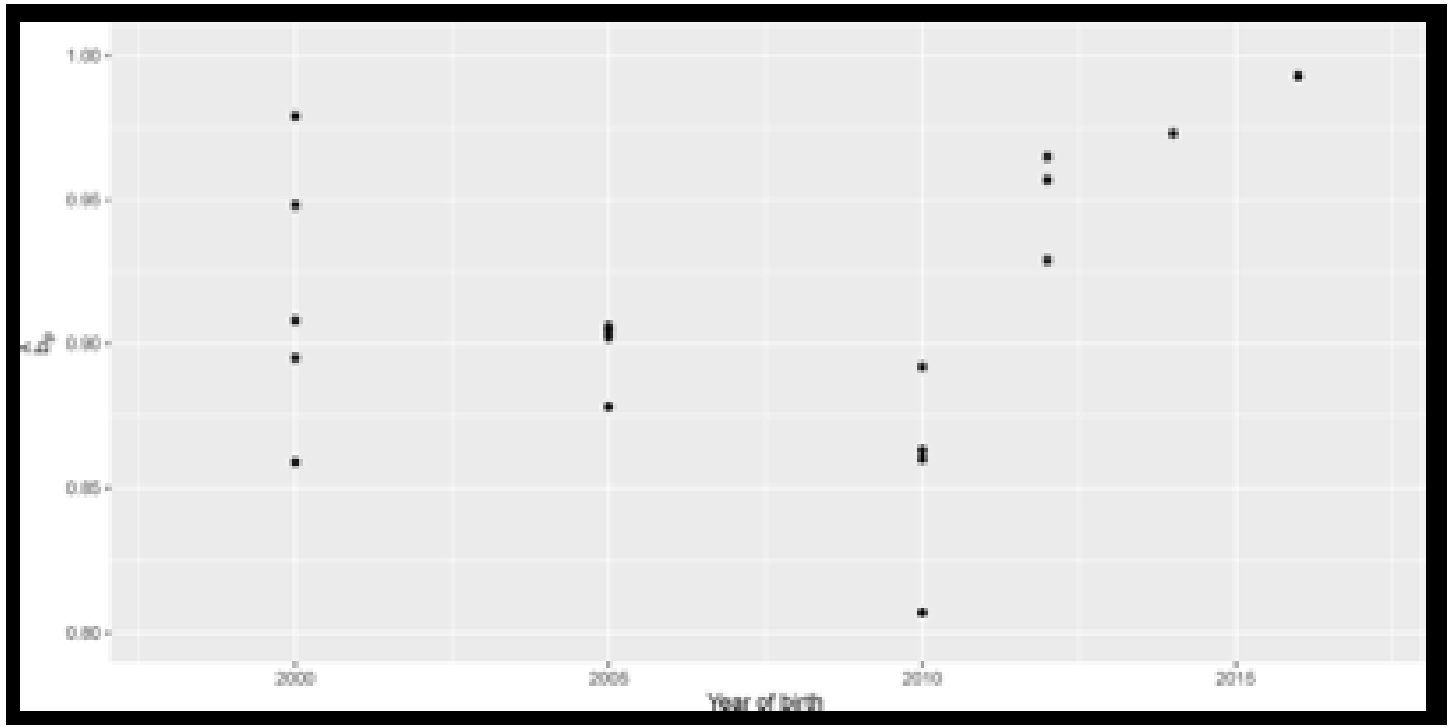
MF performs better

	EBV_w		
EBV_p	SSGBLUP-UPGA	SSGBLUP-UPGH	SSGBLUP-MF
BLUP-MF	1.32	1.29	0.98
BLUP-UPGA	1.25	1.23	0.92

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Results in Lacaune

- Slopes $\hat{b}_{w,p}$ of several cohorts of ~400 AI rams born in 2000…2016 vs. several “whole”



Results in Lacaune

- Slopes $\hat{b}_{w,p}$ of several cohorts of ~200 AI rams born in 2015…2019 (partials at 2015…2019) vs. a single “whole” at 2021

Model \ Partial →	2015	2016	2017	2018	2019
Alone	0.82	0.77	0.90	0.66	0.81
TogetherSameMF	0.90	0.87	0.93	0.83	0.89
TogetherDifferentMF	0.89	0.87	0.93	0.82	0.89
Indirect	0.08	0.28	0.09	0.06	0.09

Good models

Horrible model

Marine Wicki
PhD student
INRAE, Toulouse

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Variation in estimates

- There's a LOT of variation
- Properties of LR method (and of Interbull methods) rely on asymptotics
- strictly speaking $\widehat{\Delta_p} \to 0$ and $\hat{b}_{w,p} \to 1$ for $n \to \infty$
- so $\hat{b}_{w,p} = 1$ *either* for one year on a large sample of bulls *or* averaging small series across years
- The property of unbiasedness is a property of the prediction method as a whole (across years and animals), not of every single batch of young bulls.
- Tampering the model to avoid point biases could create further biases down the road.
- The s.e. of $\hat{b}_{w,p}$ depends on more things than the number of bulls (we can derive that)

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Sources of variation

- I see two sources
- The cohort of focal individuals may be not what we expected, i.e. they were themselves biased or not representative: $\overline{\hat{u}_p} \neq \bar{u}$
- Their further evaluations were not correct, i.e. $\hat{u}_{w(t)}$ was more "correct" in some years (t) than another years
- e.g. a value of $\hat{b}_{w,p} < 1$ may indicate overdispersion of $\hat{u}_p$ but also underdispersion of $\hat{u}_w$
- For instance this can happen if a fixed effect interacts with genetic trend in a complicated manner (I'm thinking UPGs)
- My personal suggestion (in dairy sheep) is to check models using averages across years

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

What if my model is already wrong?

- The LR theory assumes that the model is correct !!
- Can we verify if a model is correct if the model is not correct?
- Fernando Macedo explored that

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

By simulation

- Software: QMsim, Blupf90 family and our own
- 20 replicates of a “dairy” population
- 10 generations
- Two heritabilities (0.1 and 0.3)
- Three scenarios

Correct Model

Genetic evaluations performed with correct heritabilities and effects

Wrong Heritability

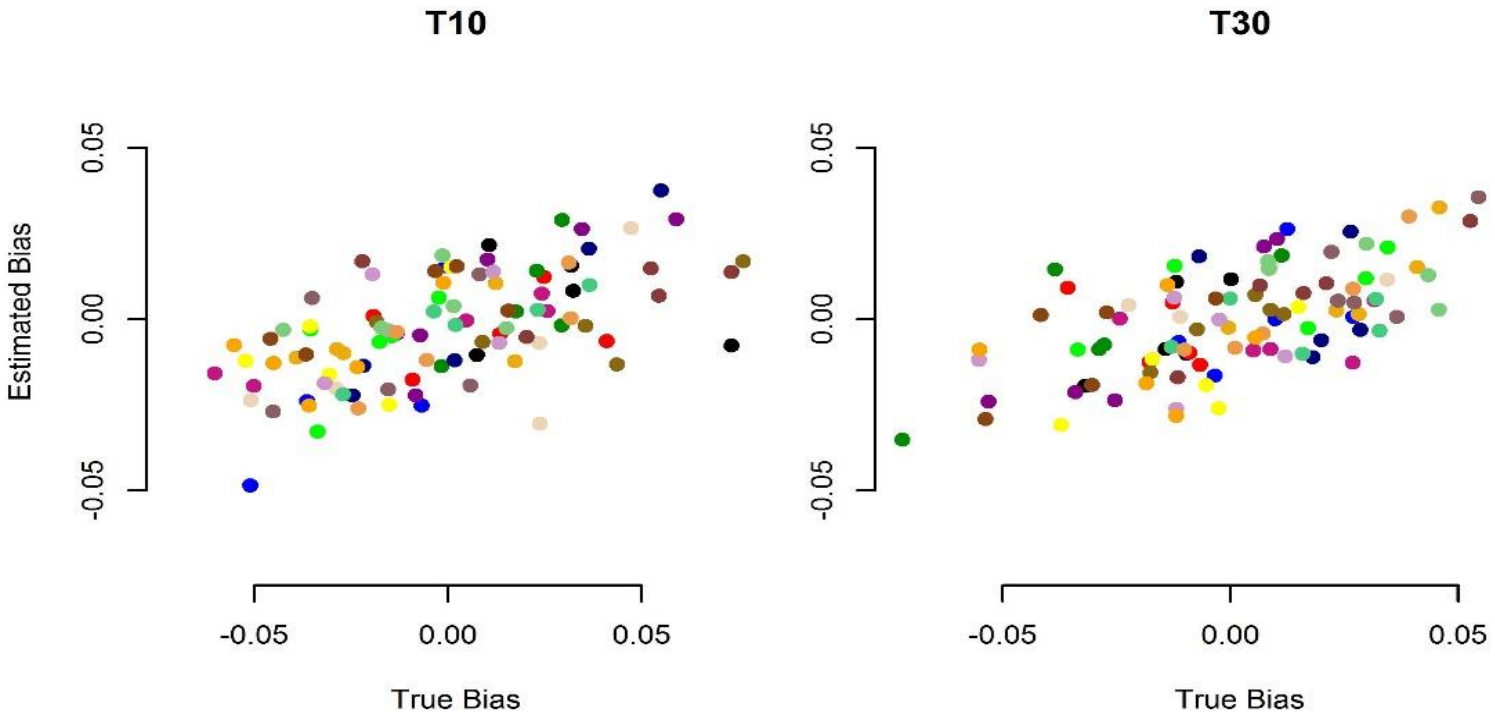
Using higher (+0.05) and lower (-0.05) heritabilities in the evaluation model

Environmental trend not (well) accounted for

Simulate a environmental trend.
Fit contemporary groups either as fixed, or as random heavily shrunken to 0 .

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

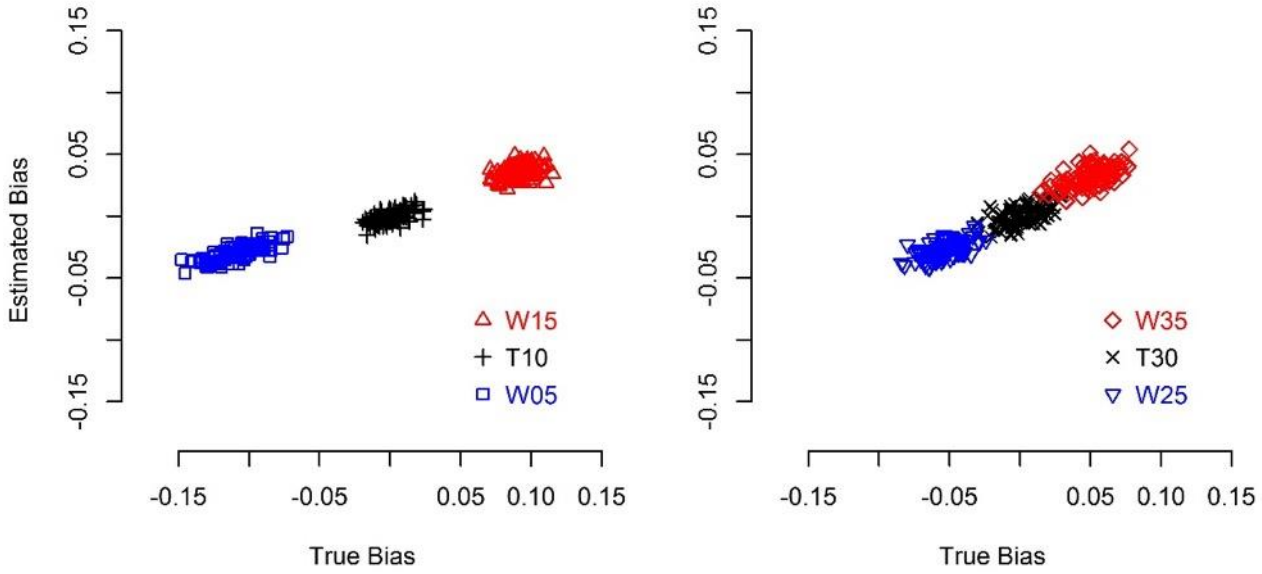
Main Results: the correct model



There was no surprise with the right model. Bias, slope, and accuracies were well estimated.

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Main Results: the wrong heritabilities



True bias was generated
LR method could estimate the good direction but not the magnitude

The slope was estimated but with low precision

Accuracies were well estimated

COUNCIL ON DAIRY CATTLE BREEDING

Main Results: the environmental trend

It was impossible to estimate the Bias, neither fitting CG as fixed nor as random effect.

The slope was poorly estimated. The estimation was better when CG were fit as fixed effects.

In general, accuracies were well estimated

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Main conclusion

The LR method can estimate the Bias, Slope, and Accuracies when the genetic evaluation model is robust, even if not perfect.

When the model is really wrong, the estimates from LR method are unreliable.

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

How does all this compare to Interbull test?

- I believe that <u>if</u> DRP/DYDs are <u>correctly</u> computed, Interbull *b1* is our "slope" $\hat{b}_{w,p}$
- Interbull doesn't check pure averages of EBVs ($\widehat{\Delta_p}$)
- In my view the logic of Interbull tests is "early=PA", "late=progeny-tested"so $Rel_{PA} \ll Rel_{progeny}$ but with genomics is this still true?

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

Take-home messages

- validation with EBVs (by method LR) allows
  - faster & automatic implementation and
  - (maybe) less approximations/incertitudes than DYDs / DRP
  - theory-based results
  - I find theory important now that we select based on early genomic predictions
- the variation in validation results needs to be considered
  - in principle we can derive theory to get e.g. the s.e. of $\widehat{\Delta}_p$, $\hat{b}_{w,p}$, etc
  - I think that checking several years is a necessity for small breeds
- I (personally) think that difference in means(GEBV) should be checked i.e. $\widehat{\Delta}_p = \overline{\widehat{\boldsymbol{u}}}_p - \overline{\widehat{\boldsymbol{u}}}_w$

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

- thank you

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

New GEBVtest Program

Pete Sullivan (Lactanet, Canada)

Background

- The Interbull software “**gebvtest.py**” is used for GEBV “Interbull validation tests”
- This Python program was modified to facilitate new and expanded validation tests of interest to the Interbull Validation Test Working Group
- the modified version is named “**gebvtest_2022C.py**”
- Backward compatibility was maintained
  - Input and output files are identical between new and old versions, if none of the optional new features are requested.

Interbull program for GEBVtest

Both the Current and New versions

Test reliability and unbiasedness of GEBV to predict future daughter performance (dEBV)

1) Reliability: $R(dEBV, GEBV_{red})^2 > R(dEBV, EBV_{red})^2$ (genomics improves R^2)
2) Unbiasedness: $b(dEBV, GEBV_{red}) = 1$ (acceptable range ~ 0.90 – 1.20)

Working Group concerns with the current tests

1) De-regression methods used to generate the dEBV can be arbitrary
2) Different bases of expression between Reduced versus Full data can directly affect the regression test, and can easily cause “FALSE FAIL” regression test results
3) EBV and dEBV carry genomic preselection bias, which means that biased GEBV look better than unbiased GEBV, and this is the opposite of what we want
4) The significance testing could be improved (e.g. with boot strapping)

Usage and New Command-line Options

gebvtest.py -h

```
gebvtest.py
usage: gebvtest.py [-h] [-v] [-m] [-M MERGEDIR] [-Z] [-C] brd pop datadir


positional arguments:
  brd                   evaluation breed code (BSW/GUE/JER/HOL/RDC/SIM)
  pop                   population code (same as country code except for
                        CHR/DEA/DFS/FRR/FRM)
  datadir               absolute or relative path to data files

optional arguments:
  -h, --help            show this help message and exit
  -v, --verbose         increase output verbosity
  -m, --mergefiles      write merged data files (for independent data checks)
  -M MERGEDIR, --mergedir MERGEDIR
                        absolute or relative path for merged data files
                        (default=DATADIR/merged)
  -Z, --no-zip          do not create a zip file (eg. for preliminary testing
                        or usage at ITBC)
  -C, --cleanup         delete all files successfully added to the zip file
```

gebvtest_2022C.py -h

```
usage: gebvtest_2022C.py [-h] [-v] [-Z] [-C]
                         [--target {DGEBV,DGPA,VFEBV,GEBV,EBV,DEBV}]
                         [--weight {ITB,LR}] [--min_byear MIN_BYEAR]
                         [--baseadj {GEBV,EBV,NONE}] [--power POWER]
                         [--baseincl BASEINCL] [--traitsincl TRAITSINCL]
                         [--outdir OUTDIR] [-m] [-M MERGEDIR]
                         brd pop datadir


positional arguments:
  brd                   evaluation breed code (BSW/GUE/JER/HOL/RDC/SIM)
  pop                   population code (same as country code except for
                        CHR/DEA/DFS/FRR/FRM)
  datadir               absolute or relative path to data files

optional arguments:
  -h, --help            show this help message and exit
  -v, --verbose         increase output verbosity
  -m, --mergefiles      write merged data files (for independent data checks)
  -M MERGEDIR, --mergedir MERGEDIR
                        absolute or relative path for merged data files
                        (default=DATADIR/merged)
  -Z, --no-zip          do not create a zip file (eg. for preliminary testing
                        or usage at ITBC)
  -C, --cleanup         delete all files successfully added to the zip file
  --target {DGEBV,DGPA,VFEBV,GEBV,EBV,DEBV}
                        validation target options are: [ DGEBV, DGPA, VFEBV,
                        GEBV, EBV, DEBV ] (default=DEBV)
  --weight {ITB,LR}     Options are: [ ITB or LR ], for the Interbull
                        weighted-regression test or Legarra-Reverter un-
                        weighted regression, respectively (default=ITB)
  --min_byear MIN_BYEAR
                        specify a minimum birth year to use instead of using
                        the value specified in the traits file
  --baseadj {GEBV,EBV,NONE}
                        evaluation variable to use for base adjustments,
                        options are: [ NONE, EBV, GEBV ] (default=NONE)
  --power POWER         specify a base for the power function weighting
                        records in base adjustments, instead of optimizing the
                        base from the data
  --baseincl BASEINCL   comma-separated lists of restrictions on bulls to
                        include for base adjustment estimates, [ min,max byr :
                        proof type list : proof status list : official Y/N ]
  --traitsincl TRAITSINCL
                        comma-separated list of traits to process
  --outdir OUTDIR       absolute or relative path to write output files
                        (default=.)
```

Enhancements to GEBVtest Program

(gebvtest_2022C.py)

1. Analyze and adjust the base of expression in Reduced data to match Full data
   - Uses a regression heavily weighted to bulls who do not add any recent data
     - Same reliability should usually mean same evaluation, if bases of expression are the same
   - Regression weights are optimized for the distribution of bull reliability changes for the trait

     --baseadj = {EBV} or {GEBV}
   - Validation tests are thus based on "*relative changes*" for bulls adding versus not adding new data
2. Allow different validation targets
   - Allowing validation targets like GEBV that are not biased by GPS (e.g. Legarra-Reverter)
   - An EBV target can also still be used, for new tests that might replace Trend Validation test III

     --target = {DEBV} or {EBV} or {DGEBV} or {GEBV}

     {DGEBV} are derived by the software using method of Van Raden (2021 Interbull bulletin)

Enhancements to GEBVtest Program

3. Allow filtering of data to use in --baseadj

   --baseincl {min_byr , max_byr : proof types : bull status : official Y|N}

   --outdir to save output files in a desired location

   - Useful for comparing test results based on different requested options

4. Command-line options to over-ride data and values read from input files

   --min_byear YYMM

   - Eliminates need to re-create several versions of input files for different tests

   --traitsincl mil,ocs,dlo

   - Useful to focus extra checks on only subsets of traits with poor test results
   - Limits --mergefiles output to only the trait subsets of interest

Enhancements to GEBVtest Program

5. Apply unweighted regression and correlation tests, instead of the standard weighted regressions used for Interbull validation tests

--weight LR

- Useful to see the impacts of using weighted regression in the Interbull tests
- In combination with --min_byear these options make it easy to see the impacts of autocorrelations on validation test results, if a weighted test is not used

Countries Reports

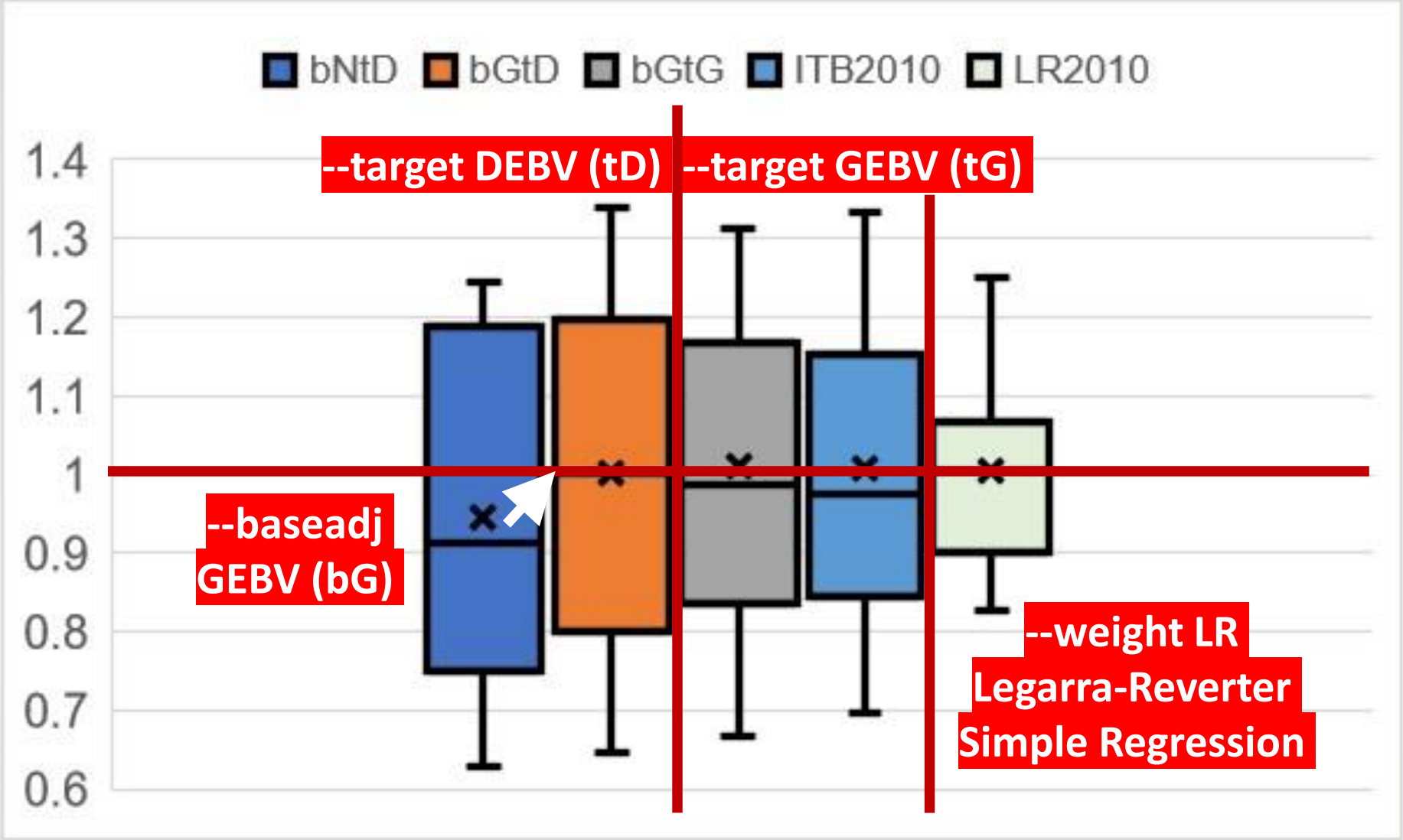
- CANADA

Applications with Canadian Data

- Full-data evaluations (EBV and GEBV) = August 2022
- Reduced-data evaluations = December 2018

- Tests applied to all traits in MACE (38)
- Summarizing results for the required test traits in GMACE (20)
  - Excludes all but 2 of the type traits

Regression tests with Canadian Data



Conclusions

- Matching the Base of Expression is important: --baseadj
- Seeking good predictions of biased future results makes no sense

  --target GEBV or DGEBV is better than the current test using --target DEBV
- We can improve on the “basic” Legarra-Reverter regression test in 2 ways:
  1. With the new --baseadj feature
  2. With weighted regressions that target animals adding data for the validation tests.
     - Reduces auto-correlations in the regression tests and increases power to detect bias
- Several new options, in combination, were designed to help focus on problem trait(s)

  --baseadj --traitsincl --mergefiles and flexible formatting now allowed for file300
  - The mergefiles include all variables for post-analyses, including extra regression variables
  - All reduced-data evaluations are base-adjusted to match the full-data base of expression
    - Average differences have expectation=0 for any group of animals

Countries Reports

- USA

Genomic validation software: USA results

Rodrigo Mota[1], Ezequiel Nicolazzi[1], and Paul VanRaden[2]

[1]CDCB - Council on Dairy Cattle Breeding, Bowie, MD, USA

[2]Animal Genomics and Improvement Laboratory, ARS, USDA, Beltsville, MD, USA

Interbull Meeting - Rome, Italy

Feb 13, 2023

CDCB

COUNCIL ON DAIRY CATTLE BREEDING

Apply new validation to 5 breeds and 7 traits

- Validation of USA genomic predictions (GPTA)
  - August 2022 official GPTA including MACE input
  - August 2018 truncated GPTA using official evaluations
- Breeds tested were HOL, JER, BSW, RDC, and GUE
- Traits tested were mil, fat, pro, scs, dlo(*), int, and mas (HO only)

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

New GEBV test software (Sullivan)

- gebvtest_2022C.py
- Minimum birth year: 2014
- Predicted deregressed dGPTA instead of dPTA
  - - - target DGEBV: uses the method of VanRaden, 2021 (Interbull Bulletin)
- - - baseadj GEBV

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Validation results: Holstein

				R^2 GEBV	R^2 PA	
Milk	3,562	1.10	0.01	74	35	Pass
Fat	3,562	1.08	0.01	77	41	Pass
Protein	3,562	1.04	0.01	74	43	Pass
SCS	3,502	1.12	0.01	69	27	Pass
Longevity (dlo)	3,330	1.01	0.01	61	30	Pass
DPR (int)	3,425	0.93	0.01	54	21	Pass
Mastitis	2,379	1.30	0.03	40	17	Fail

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Validation results: Jersey

				R^2 GEBV	R^2 PA	
Milk	648	1.06	0.03	73	49	Pass
Fat	648	1.05	0.03	63	33	Pass
Protein	648	1.05	0.03	69	45	Pass
SCS	604	1.01	0.05	45	21	Pass
Longevity (dlo)	571	0.88	0.05	36	27	Fail
DPR (int)	588	0.79	0.03	47	31	Fail

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Validation results: Brown Swiss

				R^2 GEBV	R^2 PA	
Milk	94	1.29	0.11	59	20	Fail
Fat	94	0.89	0.11	41	18	Pass
Protein	94	1.03	0.11	47	16	Pass
SCS	93	0.60	0.09	31	9	Fail
Longevity (dlo)	65	0.53	0.12	24	12	Fail
DPR (int)	88	0.64	0.16	16	27	Fail

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Validation results: Ayrshire (RDC)

				R^2 GEBV	R^2 PA	
Milk	22	0.61	0.21	35	17	Fail
Fat	22	0.91	0.20	54	33	Pass
Protein	22	0.85	0.21	50	29	Pass
SCS	21	0.79	0.30	21	26	Fail
Longevity	7	1.36	1.26	1	38	Fail
DPR (int)	22	0.60	0.52	3	1	Fail

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Validation results: Guernsey

				R^2 GEBV	R^2 PA	
Milk	19	0.92	0.22	51	36	Pass
Fat	19	0.51	0.18	32	34	Fail
Protein	19	0.77	0.25	36	26	Pass
SCS	18	1.48	0.43	42	19	Fail
Longevity	7	0.52	0.52	37	5	Pass
DPR (int)	19	0.62	0.62	18	30	Fail

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Summary

- Larger breeds and more heritable traits had more stable results
- Smaller breeds and less heritable traits are hard to validate. Tests often fail:
  - B1 more or less than expected from S.E., which may be underestimated.
  - Upper biological limit of 1.2 should allow for S.E. of B1
  - R2 of parent average may exceed GEBV with small sample sizes
- Extra regressions could help test other biases (trend, parent average, etc.)

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Acknowledgements and disclaimers

- **Pete Sullivan** of Lactanet for providing an advance copy of the validation software
- Participating **dairy producers** for supplying data
- **DHI** organizations and **DRPCs** for processing and relaying the information to the Council on Dairy Cattle Breeding (CDCB)
- **Purebred breed associations** for providing pedigree data

- Mention of trade names or commercial products is solely for the purpose of providing specific information and does not imply recommendation or endorsement by CDCB
- CDCB is an equal opportunity provider and employer

CDCB
COUNCIL ON DAIRY CATTLE BREEDING

USDA

Thank you!

Countries Reports

- DENMARK-FINLAND-SWEDEN

Test of GEBV software

Ulrik Sander Nielsen and Trine Andersen

SEGES Innovation P/S

Denmark

Interbull Technical Workshop, Rome, 14-15 February 2023

DFS HOL type traits

- **Official indexes published Nov 2022**

  Mean:100, SD:10

  App. 100 Nordic AI bulls tested per year
- **Genotype cut by birth class 2009**
- **Datasets created according to Interbull guidelines for GEBV test from 2010**
- **Reduced GEBV dataset deleting progeny of bulls born after 2014.01.01**

GEBV tests

1. Official Interbull program
2. Test program with same options as official
3. Legarra-Reverter regression
4. Legarra-Reverter regression with option –min_byear=2010
5. --target DGEBV , method of VanRaden
6. --weight=ITB –baseincl 2014,2015:11,12,10:Y
7. Official but only DFS bulls included (stature)

Results: Stature

Test	ntest	b0	b1	ncand	Exp_b1	R2
1 intb test	431	-1.74	1.00	602	0.93	80.3
2 test	431	-1.74	1.00	602	0.93	80.3
3 LR test	607	-2.74	1.01	1183	0.95	84.1
4 LR 2010	607	-2.73	1.01	2790	1.00	84.1
5 DGEBV	531	-3.65	1.02	602	0.95	84.4
6 b,2014,15	431	-1.74	1.00	602	0.94	84.1
7 DFS bulls	328	1.20	0.97	602	0.89	77.7

All tests pass: N-Y-Y-Y-Y PASS

Results: Udder support

Test	ntest	b0	b1	ncand	Exp_b1	R2
1 intb test	395	-2.74	1.02	520	0.96	80.5
2 test	395	-2.74	1.02	520	0.96	80.5
3 LR test	607	-2.74	1.02	1183	0.95	84.1
4 LR 2010	607	-2.73	1.02	2779	0.95	84.1
5 DGEBV	395	90.40	0.08	520	0.80	0.6
6 b,2014,15	395	-2.74	1.02	520	0.96	80.5

Test 1,2,3,4,6 pass: N-Y-Y-Y-Y PASS
Test 5 fails: N-N-N-N-Y FAIL

Udder support - DGEBV test

Summary statistics on candidate bulls (CB) and test bulls (TB)

Trait	Variable	N	Mean	Std	Min	Max
usu	CB EBV	520	100.468	10.029	74.56	126.37
usu	TB EBV	395	99.315	10.122	74.56	126.37
usu	TB VAL(y)	395	99.317	11.233	71.82	130.94
usu	TB GEBV(x1)	395	104.208	10.468	71.62	132.84
usu	TB EBVr(x2)	395	100.782	8.102	78.31	126.08

Details of GEBVtest calculations

usu i_est = (99.315 - 100.468) / 10.029 = -0.115

After Base adjustments:

usu p=0.947 x=-1.612 i=0.115 k=0.198 R2b=0.008 E(b1)=0.803

usu b1=0.085 se=0.054 E(b1)=0.803 t=-13.40 R2_1=0.6 R2_2=30.5

usu passes t-test=N bio-test=N b1>1=N R2-test=N b1<1.2=Y overall=FAIL

Udder support –weight=LR

```
Summary statistics on candidate bulls (CB) and test bulls (TB)
-------------------------------------------------------------
Trait Variable          N      Mean       Std      Min      Max
-------------------------------------------------------------
usu   CB EBV         1183   102.592     8.817    74.56   128.51
usu   TB EBV          607   100.825     9.419    74.56   126.37
usu   TB VAL(y)       607   105.402    11.950    63.93   137.01
usu   TB GEBV(x1)     607   106.293    10.770    71.62   138.38
usu   TB EBVr(x2)     607   102.112     8.330    78.31   126.08

Details of GEBVtest calculations
usu i_est = (100.825 - 102.592) / 8.817 = -0.200

After Base adjustments:
usu p=0.897  x=-1.263  i=0.200 k=0.293  R2b=0.882   E(b1)=0.953
usu b1=1.017  se=0.018  E(b1)=0.953  t=3.56  R2_1=84.1  R2_2=4.1
usu passes t-test=N  bio-test=Y b1>1=Y R2-test=Y b1<1.2=Y overall=PASS
```

Summary

- Without use of options, test version gives same results as official test
- Correction for base year has little effect
- LR test results in line with official test results
- Using –target DGEBV gave unexpected results for the trait, udder support

Countries Reports

- THE NETHERLANDS

Results from various validation methods

Ibrahim Jibrila, Herwin Eding

Interbull Technical Workshop

Rome, 15 Feb 2023

BETTER COWS > BETTER LIFE

Introduction

- General validation procedure
- Results
- Conclusions and discussion
- Experiences with Interbull validation software.

BETTER COWS > BETTER LIFE

General validation procedure

Routine AEU validation; for each trait in evaluation

- Full data (deregressed) GEBV used as dependent variable (GEBV or DRP)
- Reduced data (- 4 years) GEBV and PA

- Bull validations:
  – Sires with at least 20 daughters in Full, but no daughters in Reduced

- Cow validations:
  – Cows with at least one record in Full, but no observations in Reduced

BETTER COWS > BETTER LIFE

General validation procedure

Analysis

- Regress reduced run GEBV and PA on full DRP or GEBV
  - $DRP/GEBV_{ful} = a + b\ PA_{red}$
  - $DRP/GEBV_{ful} = c + d\ GEBV_{red}$ (= PA_{red} + genotype)

- Main statistics of interest:
  - Regression factor, added EDC, DGV reliability

BETTER COWS > BETTER LIFE

Testing of validation methods

- Using deregressed proofs (DRP) versus unregressed proofs (GEBV)

- Using bull validation (classical) versus cow validation.
  – Could possibly alleviate lack of data issues

- Results shown for illustrative traits
  – Udderhealth
  – Calving ease
  – Ketosis
  – Reproductive disorders

Results: Regression factors

Trait name	h2	N buls	N cows	Regression coeff			
				bul drp	bull gebv	cow drp	cow gebv
subclinical mastitis	0,056	763	89776	1,03	0,99	1,12	1,02
clinical mastitis	0,062	775	14698	0,89	0,86	1,08	0,96
udder health index	0,090	775	90196	0,96	0,92	1,11	0,99
maternal stillbirth heifers	0,084	571	81396	1,06	1,03	1,21	1,04
maternal stillbirth cow	0,005	485	71515	1,12	1,03	1,33	1,06
direct stillbirth heifers	0,041	444	38741	0,94	0,90	1,20	0,93
direct stillbirth cows	0,006	571	66131	0,96	0,88	1,22	0,94
milk fever parity	0,035	51	3565	0,59	0,61	1,09	0,94
clinical ketosis overall	0,096	115	8619	0,62	0,60	0,87	0,75
retained placenta	0,064	272	17672	0,69	0,65	0,90	0,78
endometritis	0,060	256	17519	0,67	0,65	0,93	0,79
metritis	0,049	224	14543	0,61	0,60	1,01	0,81
cystic ovaries	0,029	210	12186	0,85	0,79	1,00	0,85
anoestrus	0,034	293	18325	0,70	0,67	0,89	0,77
index reproduction disorders	0,123	337	22646	0,71	0,65	0,93	0,76

Results: Added EDC

Trait name	h2	N buls	N cows	EDC added			
				bul drp	bull gebv	cow drp	cow gebv
subclinical mastitis	0,056	763	89776	145,4	160,8	374,2	517,0
clinical mastitis	0,062	775	14698	115,4	134,4	275,4	389,2
udder health index	0,090	775	90196	87,2	97,6	216,9	302,3
maternal stillbirth heifers	0,084	571	81396	81,1	86,8	154,2	210,6
maternal stillbirth cow	0,005	485	71515	803,7	967,5	1081,2	1597,3
direct stillbirth heifers	0,041	444	38741	48,9	53,9	88,9	156,4
direct stillbirth cows	0,006	571	66131	318,2	373,2	576,3	1052,8
milk fever parity	0,035	51	3565	14,4	24,5	33,7	64,0
clinical ketosis overall	0,096	115	8619	6,3	8,1	6,4	19,9
retained placenta	0,064	272	17672	7,1	5,0	32,1	47,1
endometritis	0,060	256	17519	8,5	8,1	33,3	46,8
metritis	0,049	224	14543	8,3	7,6	32,1	44,7
cystic ovaries	0,029	210	12186	27,5	28,0	53,6	74,4
anoestrus	0,034	293	18325	16,6	16,6	53,3	81,3
index reproduction disorders	0,123	337	22646	4,5	4,0	14,4	20,1

Conclusions

- DRP validation seems more consistent then GEBV validation
- Bull validation seem more stable then cow validation
  - Unless lack of data is an issue. Cow validations can be used when lack of validation bulls occurs.
- Added EDC (diff in regression R^2) higher in GEBV validation then in DRP validation
- Added EDC in cow validations higher then in bull validation
- Leads to inflated estimates of mean DGV reliability
  - Validations of GEBV on bull DRP seems preferable.
  - In a pinch cow DRP validations can be done.

BETTER COWS > BETTER LIFE

AEU versus ITB validations

New validation software made available

Limited testing on a few number of traits

- Compare validation on
  - EBV (Full conventional)
  - GEBV (Full genomic)

AEU vs ITB validation: EBV validation

Trait name	B1				Diff _R2			
	aeu_debv	aeu_ebv	itb_debv	itb_ebv	aeu_debv	aeu_ebv	itb_debv	itb_ebv
Milking speed	0,74	0,70	0,89	0,74	0,35	0,35	0,18	0,07
Temperament	0,61	0,62	1,16	0,99	0,07	0,03	0,03	-0,02
Direct stillbirth	0,55	0,51	0,43	0,44	0,04	0,02	0,03	0,01
Maternal stillbirth	0,64	0,65	0,71	0,77	0,08	0,06	0,01	-0,02
Clinical mastitis	0,56	0,56	1,48	1,55	0,03	0,12	0,07	0,06

AEU vs ITB validation: GEBV validation

Trait name	B1				Diff _R2			
	aeu_dgebv	aeu_gebv	itb_dgebv	itb_gebv	aeu_dgebv	aeu_gebv	itb_dgebv	itb_gebv
Milking speed	1,07	1,03	1,10	1,05	0,69	0,71	0,27	0,27
Temperament	0,92	0,89	0,80	0,90	0,52	0,56	0,16	0,23
Direct stillbirth	0,93	0,89	0,70	0,84	0,33	0,35	0,09	0,19
Maternal stillbirth	0,99	0,96	0,76	0,85	0,53	0,55	0,12	0,17
Clinical mastitis	0,85	0,82	0,79	0,85	0,58	0,62	0,18	0,28

CRV
BETTER COWS > BETTER LIFE

AEU vs ITB validation

- GEBV validation shows better agreement (AEU, ITB)
  - Higher succes rate
  - Coventional EBV biased?
    Does not account for genomics in PA => less predictability
    False FAILs

- Validation on GEBV seems to be preferable
  - For both AEU and ITB validations

- Testing of new Interbull software continues...

BETTER COWS > BETTER LIFE

BETTER COWS > BETTER LIFE

Countries Reports

- GERMANY

IT-Solutions for
Animal Production

Validating German Holstein single-step evaluations for test-day traits using Interbull's new GEBVtest software

Zengting Liu and Hatem Alkhoder
IT Solutions for Animal Production (**vit**), Germany

Scenarios for testing the new software for GEBV test

- **Dependent variable**
  - Full evaluation GEBV
  - Deregressed GEBV (VanRaden, 2021)
  - Deregressed EBV (DRP) from the current conventional evaluations
    - Deregressed MACE EBV for bulls (MACE_DRP)
    - Deregressed national EBV for cows with national phenotype data only (NAT_DRP)

- **Validation animals**
  - Genotyped bulls having daughters in DEU and EDC ≥ 20 (1,655 Holstein bulls)
  - Genotyped domestic cows with test-day records (180,389 Holstein cows)
    - Own test modification of the software for validation cows: removal of the minimum EDC 20

- Full and truncated evaluations both adjusted for the base population average
  - Using the same cow base population
  - Option of *no base adjustment* for the GEBVtest software

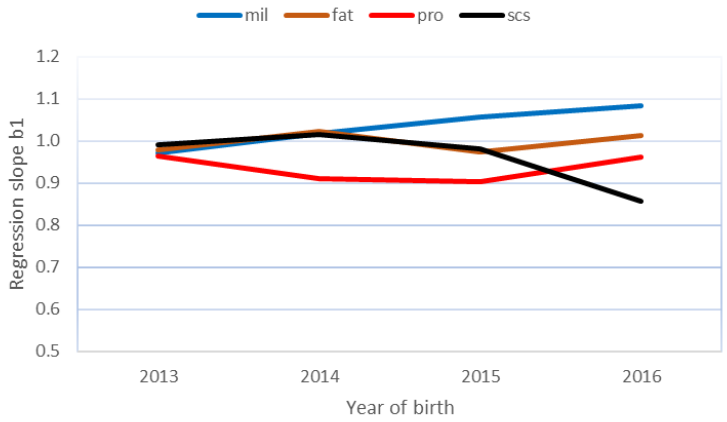
Validation results using full evaluation GEBV as dependent variable (I)

- **Validation bulls**

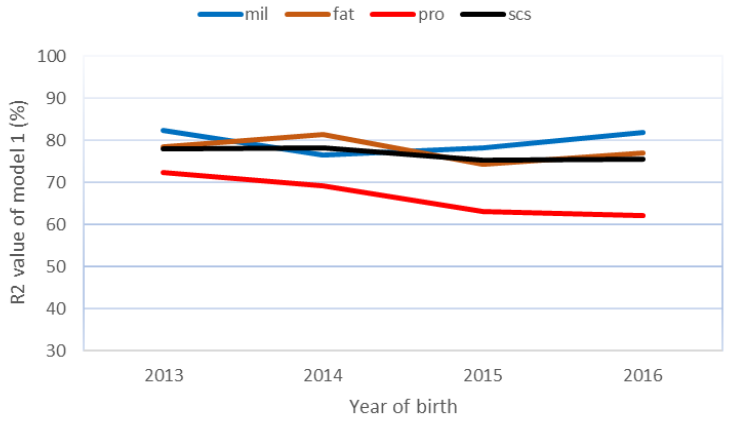
- Model 1: $GEBV_{full} = b_0 + b_1 * GEBV_{trunc}$
  - Weighted LR regression
- Model 2: $GEBV_{full} = b_0 + b_1 * EBV_{trunc}$

Trait	Model 1		M1-M2	
	b_1	R^2	ΔR^2	Pass
Milk yield	1.01	0.80	0.62	Y
Fat yield	1.00	0.80	0.50	Y
Protein yield	0.95	0.71	0.47	Y
SCS	0.99	0.78	0.54	Y

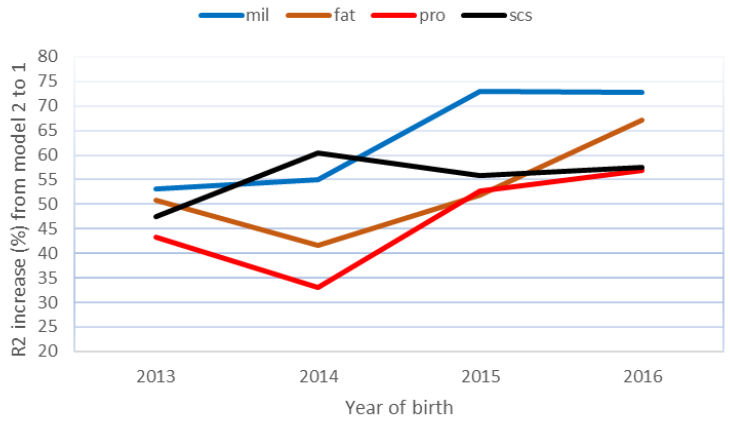
b_1 of Model 1



R^2 of Model 1



ΔR^2: Model 1-2



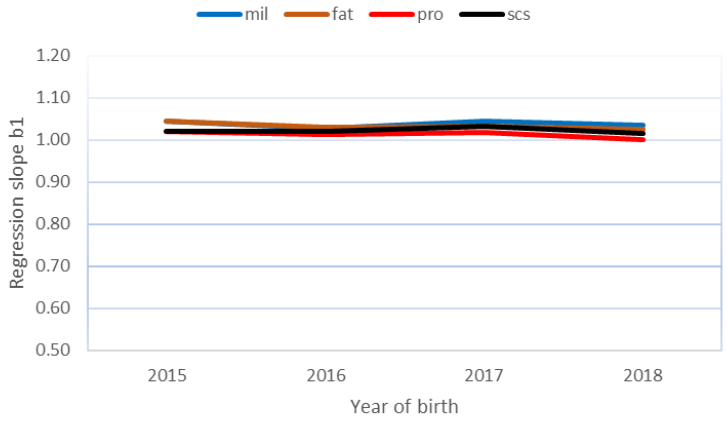
Validation results using full evaluation GEBV as dependent variable (II)

- **Validation cows**

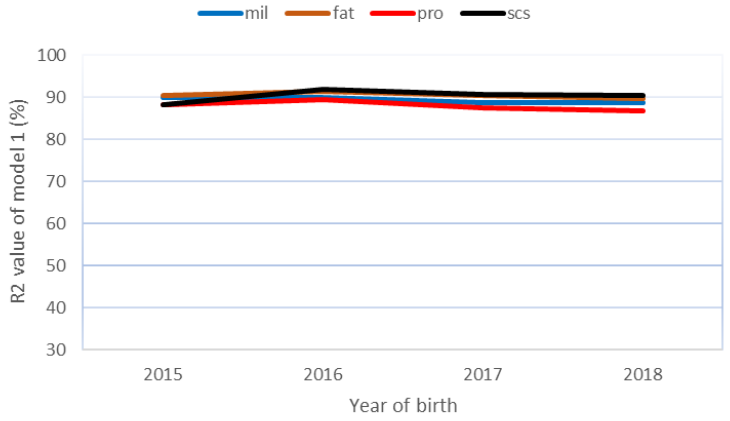
- Model 1: $GEBV_{full} = b_0 + b_1 * GEBV_{trunc}$
  - Weighted LR Regression
- Model 2: $GEBV_{full} = b_0 + b_1 * EBV_{trunc}$

Trait	Model 1				M1-M2		
	b_1	bulls	R^2	bulls	ΔR^2	bulls	Pass
Milk yield	1.03	1.01	0.89	0.80	0.59	0.62	Y
Fat yield	1.03	1.00	0.91	0.80	0.55	0.50	Y
Protein yield	1.01	0.95	0.88	0.71	0.48	0.47	Y
SCS	1.02	0.99	0.91	0.78	0.66	0.54	Y

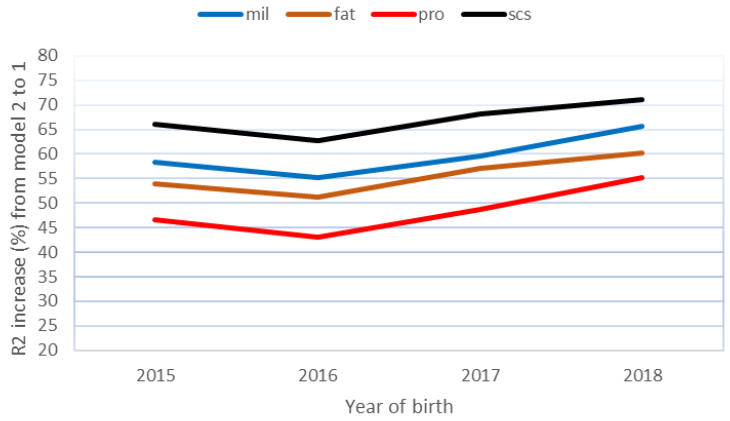
b_1 of Model 1



R^2 of Model 1



ΔR^2: Model 1-2



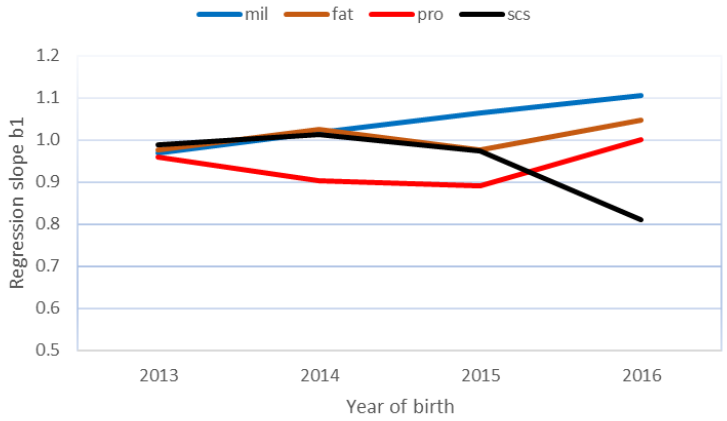
Validation results using VanRaden deregressed GEBV as dependent variable (I)

- **Validation bulls**

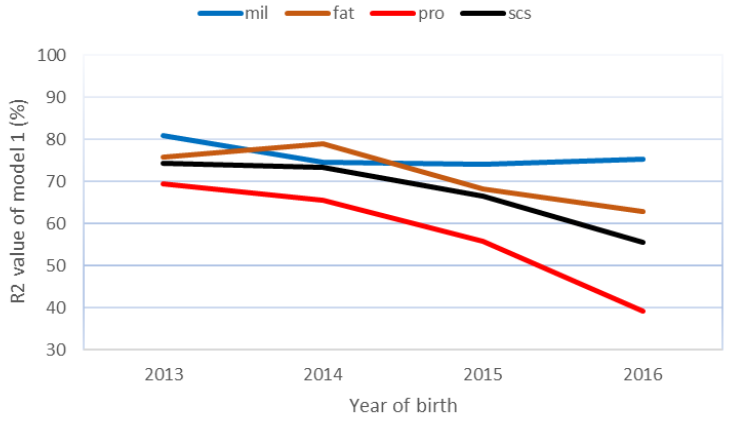
- Model 1: $DGEBV_{full} = b_0 + b_1 * GEBV_{trunc}$
- Model 2: $DGEBV_{full} = b_0 + b_1 * EBV_{trunc}$

Trait	Model 1				M1-M2		
	b_1	GEBV	R^2	GEBV	ΔR^2	GEBV	Pass
Milk yield	1.01	1.01	0.77	0.80	0.60	0.62	Y
Fat yield	1.00	1.00	0.75	0.80	0.48	0.50	Y
Protein yield	0.94	0.95	0.64	0.71	0.43	0.47	Y
SCS	0.98	0.99	0.71	0.78	0.50	0.54	Y

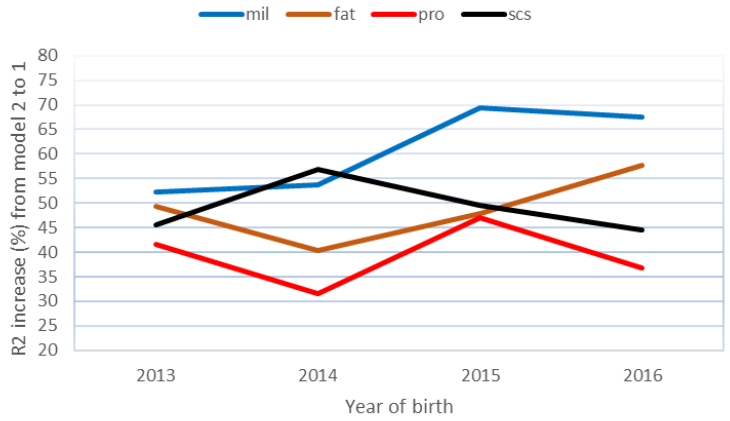
b_1 of Model 1



R^2 of Model 1



ΔR^2: Model 1-2



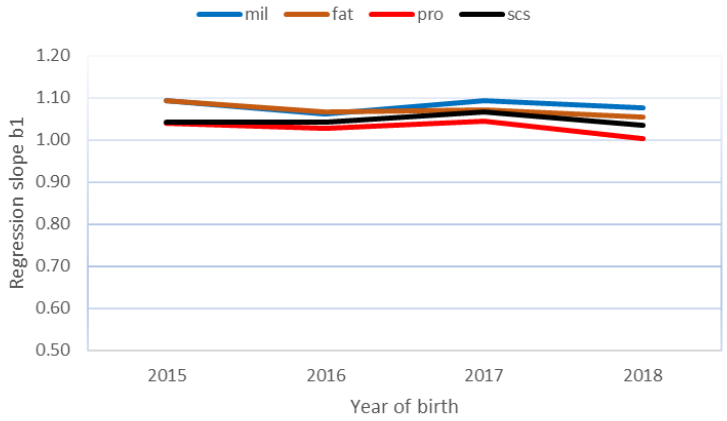
Validation results using VanRaden deregressed GEBV as dependent variable (II)

- **Validation cows**

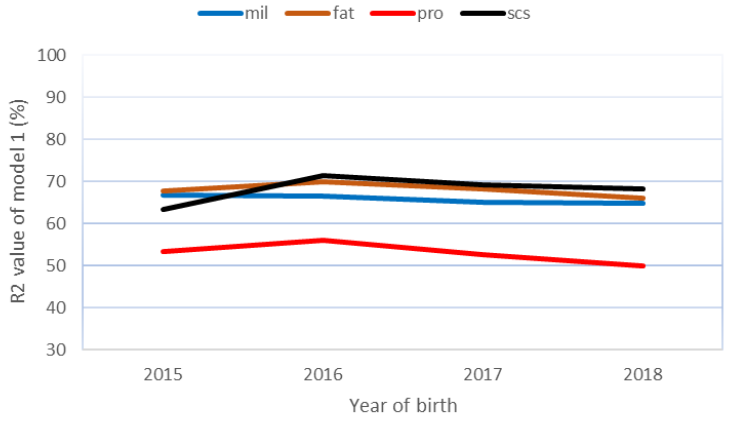
- Model 1: $DGEBV_{full} = b_0 + b_1 * GEBV_{trunc}$
- Model 2: $DGEBV_{full} = b_0 + b_1 * EBV_{trunc}$

Trait	Model 1				M1-M2		
	b_1	GEBV	R^2	GEBV	ΔR^2	GEBV	Pass
Milk yield	1.07	1.03	0.66	0.89	0.46	0.59	Y
Fat yield	1.06	1.03	0.68	0.91	0.45	0.55	Y
Protein yield	1.02	1.01	0.54	0.88	0.30	0.48	Y
SCS	1.05	1.02	0.69	0.91	0.52	0.66	Y

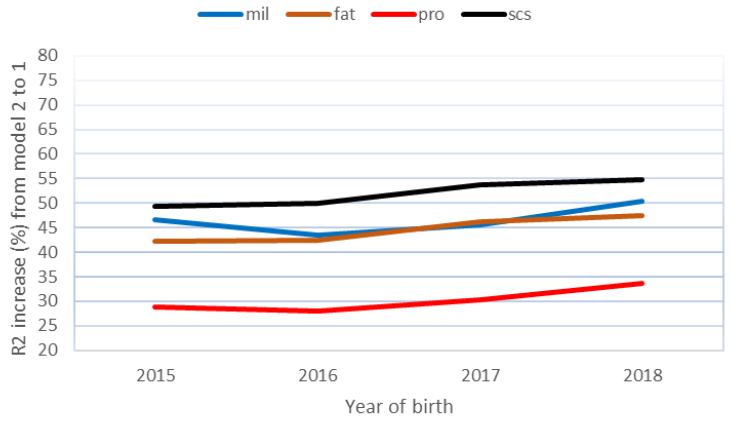
b_1 of Model 1



R^2 of Model 1



ΔR^2: Model 1-2



Summary and conclusions (I)

- The new GEBVtest software successfully tested with DEU single-step model for test-day traits using
  - Three types of dependent variables, and
  - Two groups of validation animals
  - And also for the current 2-step genomic model (DEU conformation traits)
- For all tested scenarios, b_1, R^2 and ΔR^2 seem to meet expectations
  - Validation cows (low reliability animals) and bulls (high reliability) behaved differently
  - Validation results vary across subgroups of validation animals, e.g. born in different years
- Deregressed GEBV / EBV resulted in lower model R^2 values than GEBV
  - Not directly comparable between the two dependent variables DGEBV and GEBV
  - Regression slope b_1 deviated slightly more from 1
- For high reliability validation animals, e.g. national bulls, dependent variables GEBV and DGEBV led to more similar R^2 values than for low reliability validation animals, e.g. cows
  - Depending on the contribution of own phenotype data to the total reliability of animal

Summary and conclusions (II)

- Issues concerning the genomic validation method
  - GEBV as dependent variable for low-reliability validation animals → extremely high R^2 value
    - Own phenotype data contributing less to own GEBV than bulls with many daughters
  - Validated genomic reliabilities as input data for VanRaden's DGEBV, e.g. using Interbull GREL method
  - Deregressed GEBV as dependent variable are more desirable than GEBV
    - Alternative way of computing deregressed GEBV for single-step model (Liu and Masuda, 2021)
      - Using genomic and pedigree relationship matrices
      - Iterative procedure of solving deregressed GEBV
      - Independent of the truncated genomic evaluation
  - 4-year truncation of phenotype data may be revised for more realistic forward prediction
    - A short history of large-scale cow genotyping in most countries
  - Requirement of special conventional evaluations (Cf and Cr) in the era of single-step evaluation
    - Testing $\Delta R^2 > 0$ still necessary after ~15 years of genomic selection?

Thanks for your attention!

IT-Solutions for Animal Production

Acknowledgement: Dr. S. Rensing and Dr. J. Heise for discussion

GENERAL DISCUSSION