



➤ An approach to reduce computing time in multi-trait single-step evaluations

L.-H. Maugan. T. Tribout. V. Ducrocq

Université Paris-Saclay. INRAE. AgroParisTech. GABI. 78350. Jouy-en-Josas. France

> Context

Single-Step evaluation

→ SS evaluation is more and more considered as the standard for genetic evaluation when selection strongly relies on genomic information of young animals.

Requires new software development to be feasible on large populations

But complex and time consuming, in particular due to slow convergence → restriction to (often simplified) univariate models

Drawbacks: loss of information from correlated traits and/or simplified assumptions (e.g. neglecting heterogeneity of residual variance)

Multivariate SS software exist but usually assume the same underlying model for all traits → loss of flexibility

> Context



A potential alternative: Combined SS evaluation

= **Perform a multivariate SS evaluation ... in two steps**

- Run univariate evaluations adapted to the trait (i.e. with an accurate modelling) and derive a corrected phenotype (e.g. corrected for fixed effects and heterogeneity of variances) with an associated weight.
- Run a multivariate SS evaluation of corrected+weighted performances with a simplified model

Expected benefits: improve accuracy for less heritable traits and hopefully decrease computing time

Note: this is the approach used in France for GENETIC evaluations since 2001 to derive Total Merit Index with EBV for all traits for all animals.

It is also used to improve GENETIC evaluation for Functional Longevity adding information from traits correlated to Functional longevity



➤ Material and Method



As a first example, we started with a simple situation: a group of correlated type traits in Montbéliarde breed with traits collected together and described by a same model

		HS	CW	CD	BD	RL	WH	TW	RA
Height at sacrum	HS	0.63							
Chest width	CW	0.38	0.33						
Chest depth	CD	0.66	0.51	0.46					
Body depth	BD	0.62	0.53	0.82	0.42				
Rump length	RL	0.75	0.52	0.70	0.59	0.40			
Width at hips	WH	0.59	0.59	0.64	0.56	0.67	0.39		
Width at thurl	TW	0.60	0.63	0.61	0.59	0.62	0.79	0.38	
Rump Angle	RA	-0.12	0.10	0.09	0.05	0.03	0.23	0.35	0.36

Heritability (diagonal) and genetic correlation (off-diagonal) between traits

Note that the current official GENETIC evaluation takes into account heterogeneous residual variances associated to year-classifier. This heterogeneity of variance is not (not yet) considered here.



➤ Material and Method

Models

1) Standard Multitrait Single-Step approach

$$y_{kj} = \sum_1^3 b_k + a_j + e_{kj}$$

y_{kj} : animal performance

b_k : fixed effects (calving age, lactation stage and herd x day x classifier)

a_j : additive genetic effect of animal j

e_{kj} : residual

HSSGBLUP software developed at INRAE was used for this study.

Every 200 SS iterations, (G)EBVs are compared to final results.

Convergence at iteration i was measured through: mean and standard deviation of the absolute difference between current solutions and final results and correlation with final results.

2) Combined Single-Step approach

Step 1 : Univariate model:

$$y_{kj} = \sum_1^3 b_k + a_j + e_{kj}$$

Step 2 : Multitrait model:

$$Y_{kj_corrected \text{ for fixed effects}} = \mu_{an} + a_j + e_{kj}$$

$Y_{kj_corrected \text{ for fixed effects}}$: animal performance corrected for fixed effects of step 1

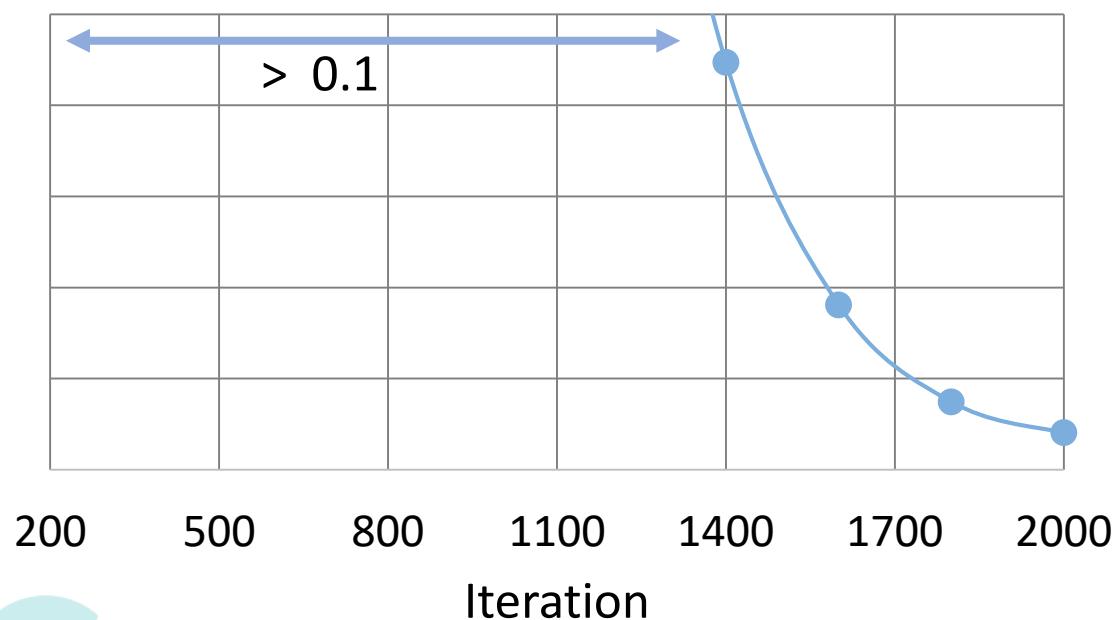
μ_{year} : a year effect

a_j : additive genetic effect of animal j

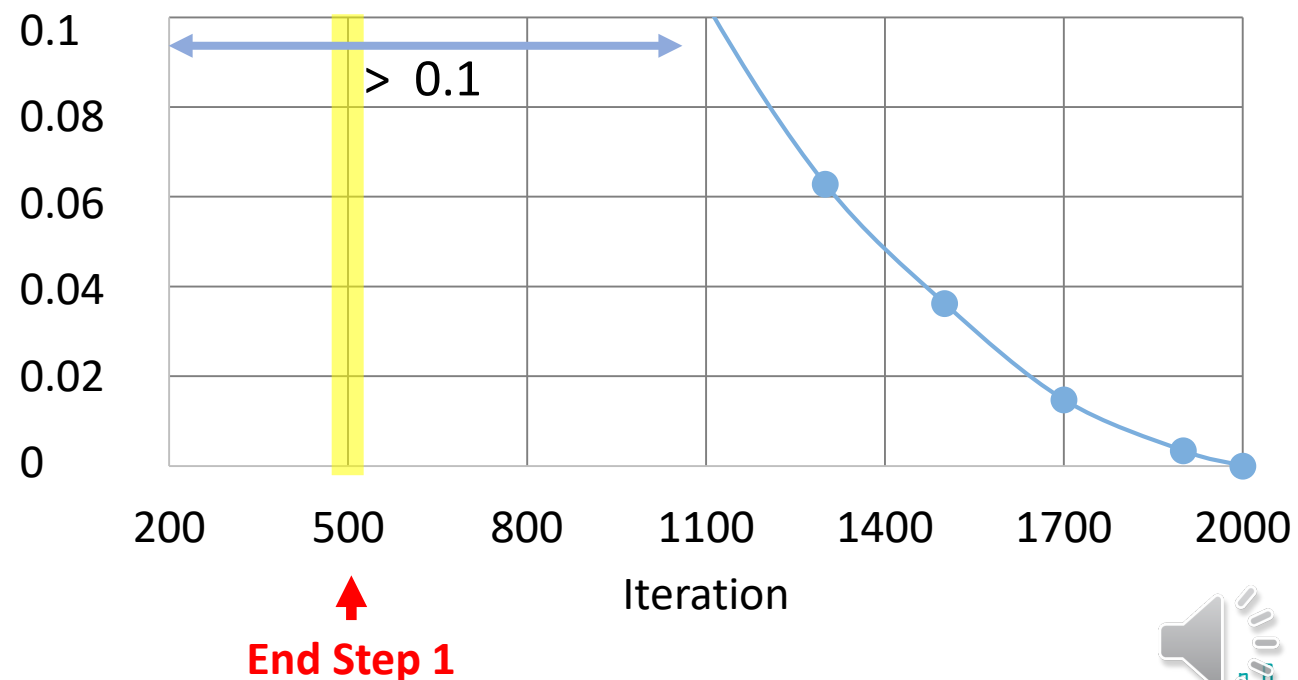
e_{kj} : residual

➤ Results for Rump Length

Standard Deviation of the |difference| between GEBV at iteration n and final (=iteration 3000) GEBV in standard **Multitrait Single-Step** approach

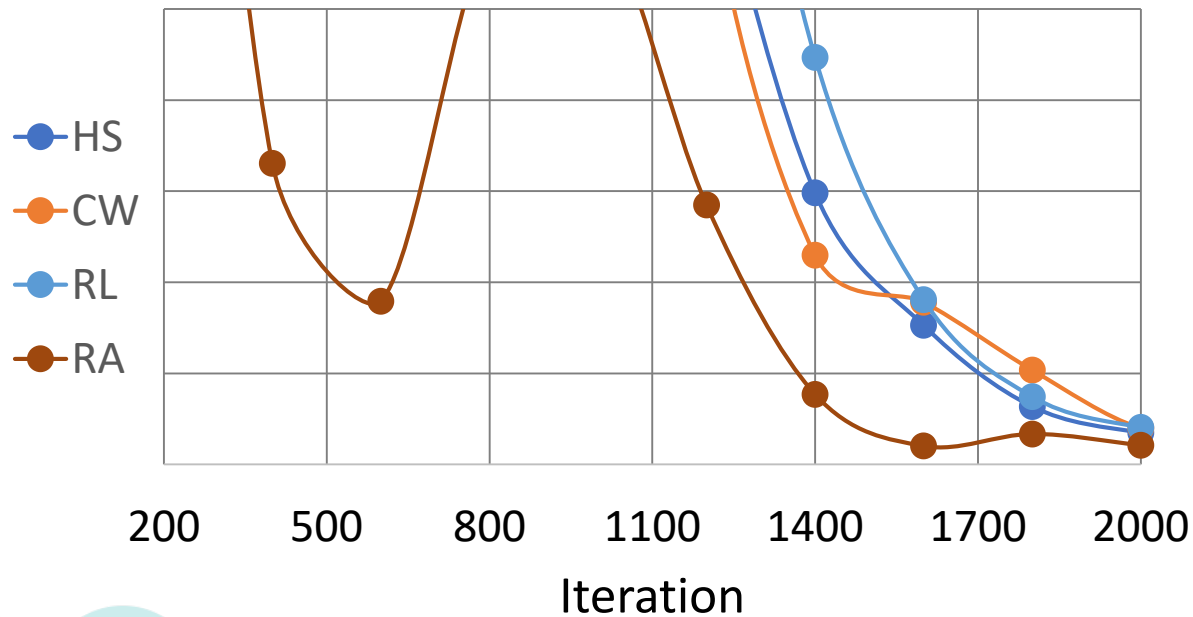


Standard Deviation of the |difference| between GEBV at iteration n and final GEBV in **Combined Single-Step** approach

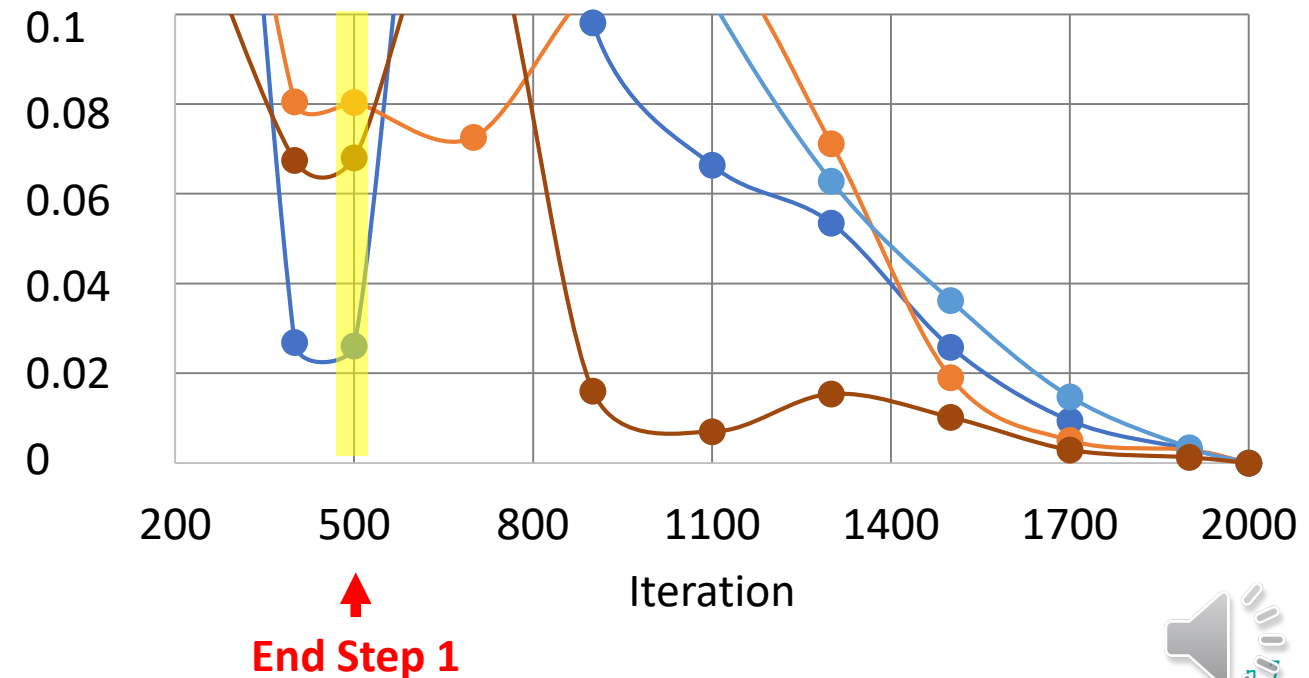


➤ Results

Standard Deviation of the |difference| between GEBV at iteration n and final (=iteration 3000) GEBV in standard **Multitrait** Single-Step approach



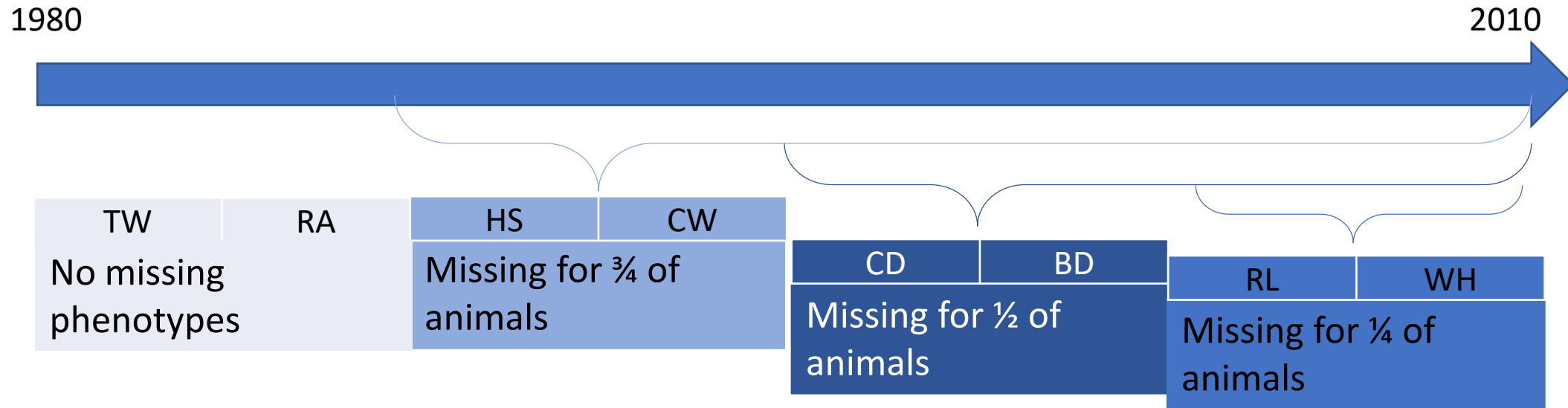
Standard Deviation of the |difference| between GEBV at iteration n and final GEBV in **Combined** Single-Step approach



> Other application



We also considered a situation with missing records for an increasing number of traits over time



The structure and the proportion of missing records have a negative impact on convergence. This impact is more acute for multitrait SS models than for Combined SS models.

(We also considered the opposite situation with fewer traits at the beginning)

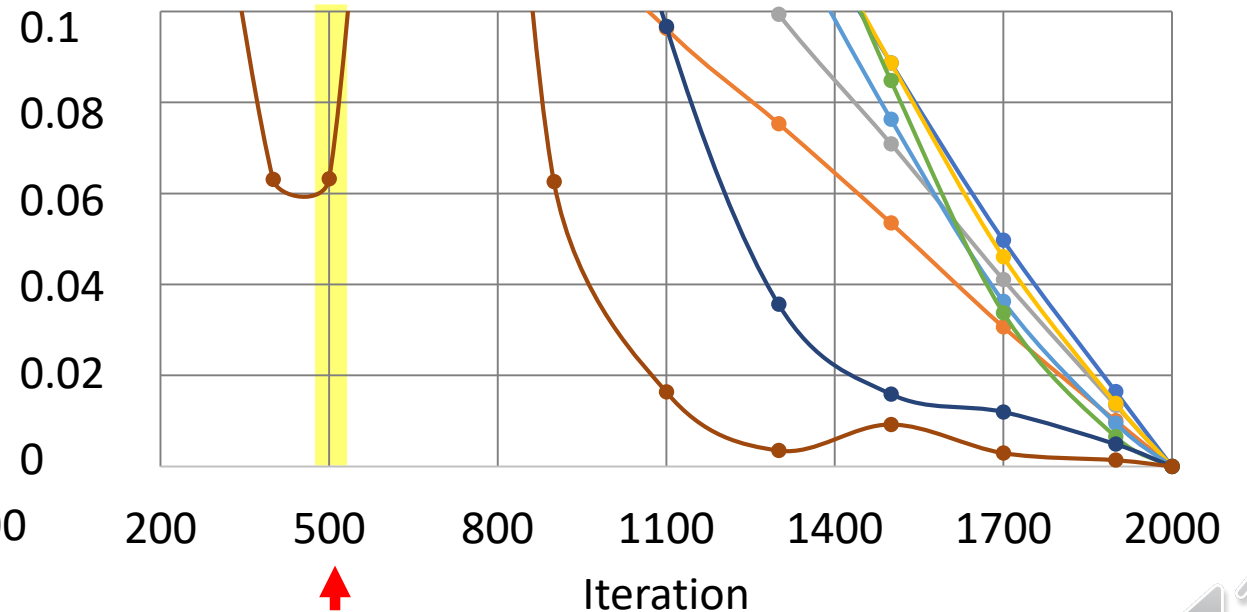
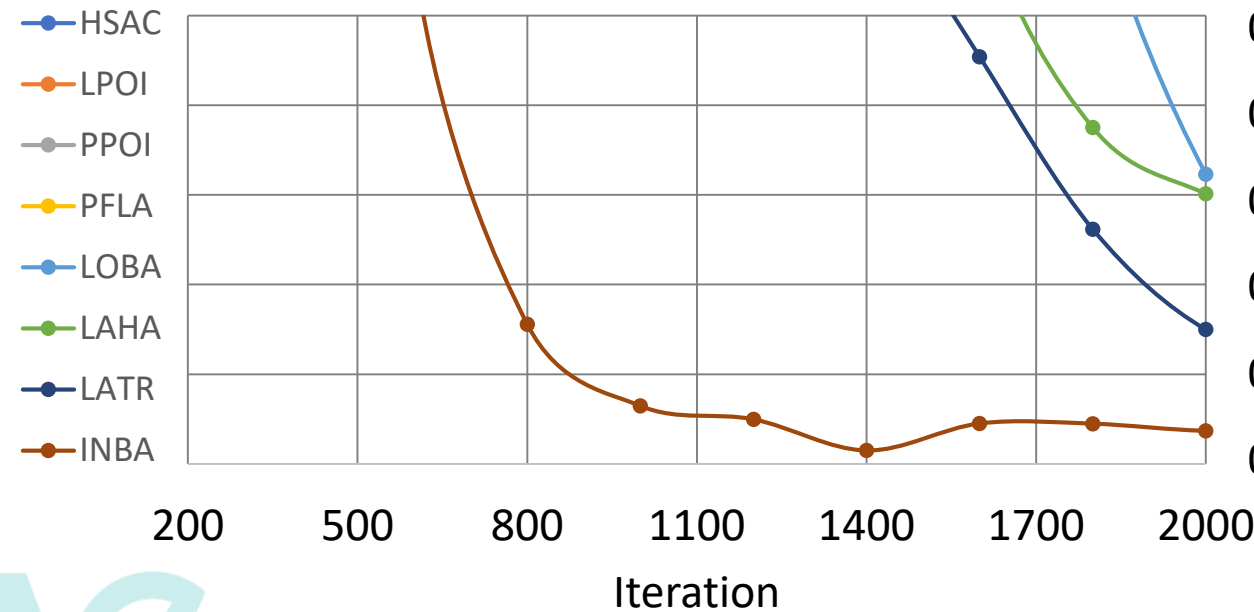


➤ Other application

Mean of the |difference| between GEBV at iteration n and final (=iteration 3000) GEBV in standard **Multitrait** Single-Step approach

Mean of the |difference| between GEBV at iteration n and final GEBV in **Combined** Single-Step approach

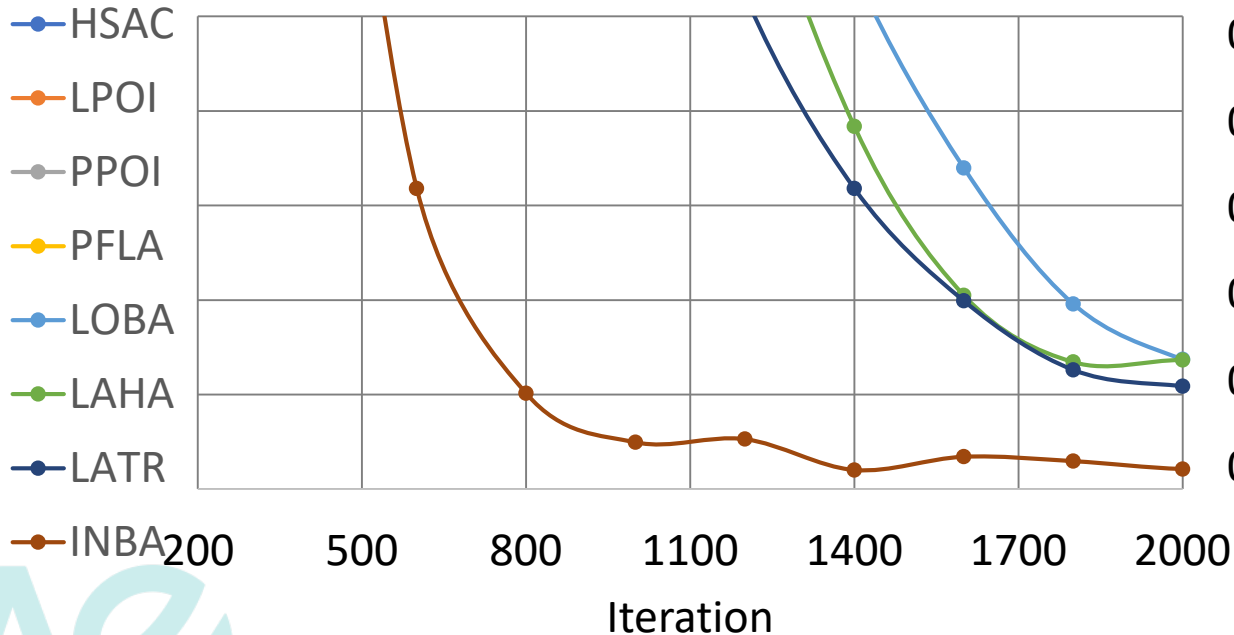
↔ > 0.1 for the 4 other traits



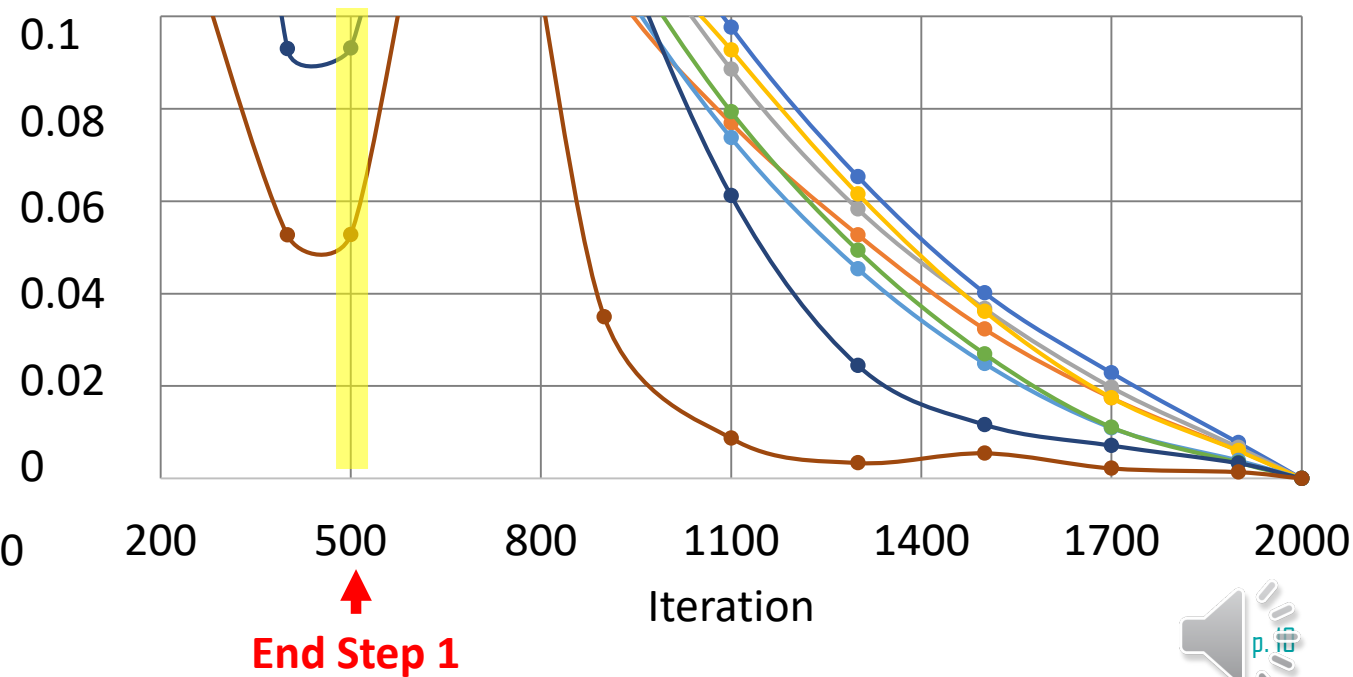
➤ Results

Standard Deviation of the |difference| between GEBV at iteration n and final (=iteration 3000) GEBV in standard **Multitrait Single-Step** approach

↔ > 0.1 for the 4 other traits



Standard Deviation of the |difference| between GEBV at iteration n and final GEBV in **Combined Single-Step** approach

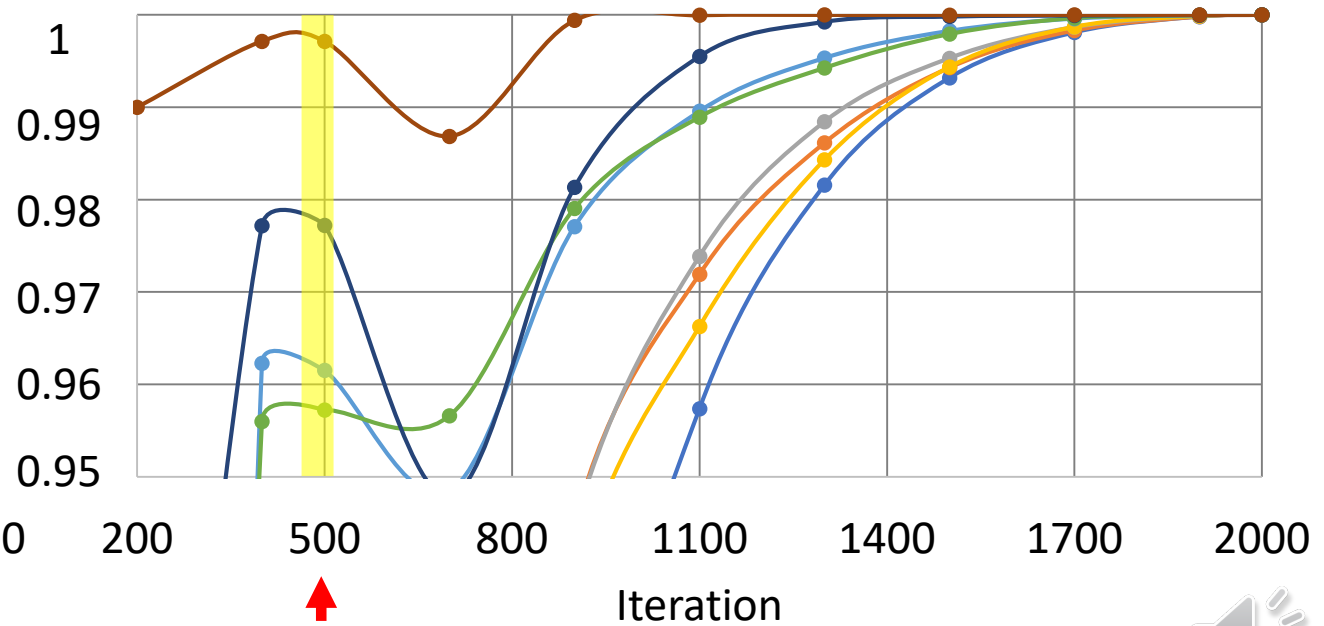
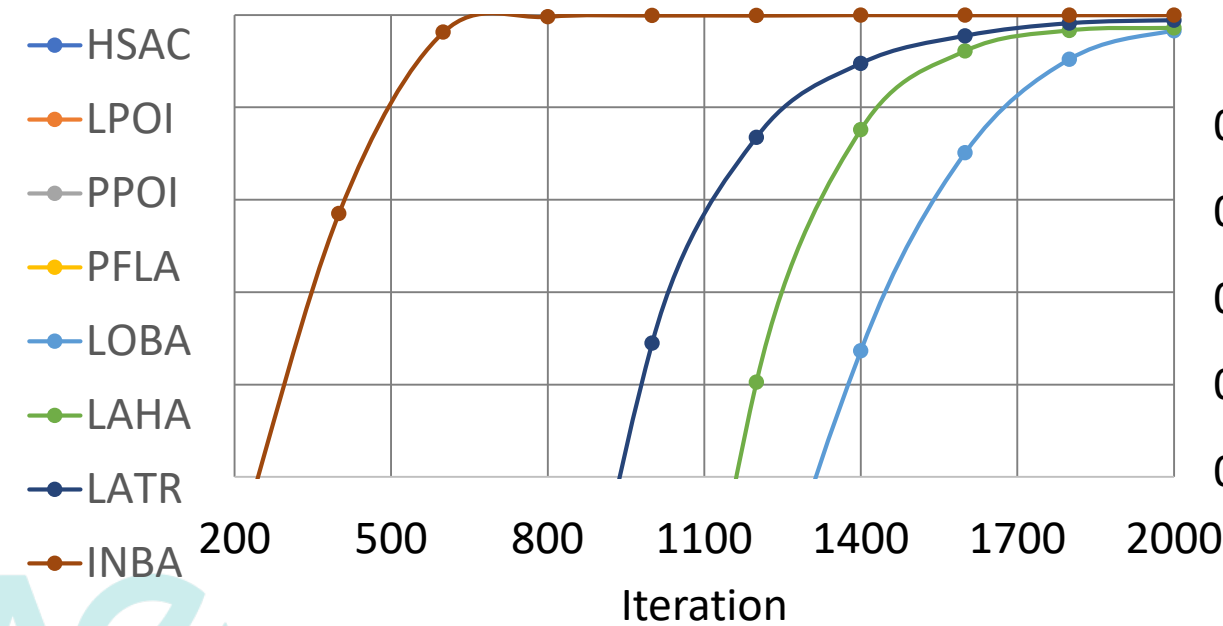


➤ Results

Correlation between GEBV at iteration n and final (=iteration 3000) GEBV in standard **Multitrait** Single-Step approach

Correlation between GEBV at iteration n and final GEBV in **Combined** Single-Step approach

← < 0.95 for the 4 other traits →



> Conclusion



Combined Single-Step approach converges faster than a standard Multitrait approach (fewer iterations)

→ could decrease running time by reducing number of iterations **especially when we have missing phenotypes**

Some differences between traits according to their heritability and genetic correlations, as in standard Multitrait approach

Limit of this example:

Simple group of rather highly correlated traits



➤ Conclusion

Next steps

- Illustrate the benefits of the inclusion of heterogenous variances due to classifiers
- Consider groups of traits described by different models (e.g. functional longevity + correlated traits)

Main message:

Implementing a Single-Step evaluation in two steps increases modelling flexibility

This work will continue during my PhD thesis financed by the UniGeno CASDAR project



➤ Thanks for watching



If you want to ask question, I'll be on the «New developments in Single Step genomic evaluations and validation methods » session on Monday, April 26 at 16:05



(Sorry for the English mistakes!)

26th April / Maugan

