

HSSGBLUP: a single-step SNP-BLUP genomic evaluation software adapted to large livestock populations.

Thierry Tribout, Vincent Ducrocq, D Boichard

▶ To cite this version:

Thierry Tribout, Vincent Ducrocq, D Boichard. HSSGBLUP: a single-step SNP-BLUP genomic evaluation software adapted to large livestock populations.. International Congress of Quantitative Genetics, Nov 2020, Melbourne, Australia. . hal-03155472

HAL Id: hal-03155472 https://hal.inrae.fr/hal-03155472

Submitted on 1 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Liberté Égalité Fraternité



Context

In France, current dairy and beef cattle genomic evaluations are based on multi-step approaches. Preselection of genotyped animals generates biased estimated breeding values and genetic trends.

Single Step GBLUP evaluations are being implemented to solve this issue, with the development of a software fitting the French bovine evaluations requirements:

- Large populations (up to 20 million animals)
- Hundreds of thousands of informative genotyped animals
- Genetic Groups
- Multiple traits evaluations, possibly with maternal genetic effects and heterogeneous variances
- Inclusion of effects of QTL or causal variants
- Inclusion of foreign phenotypic information for international populations (Holstein, BSW)

INRAE develops a software covering these features: *HSSGBLUP*. The main strategies adopted are presented here.

Current status & Perspectives

The software is completed. Optimizations (computing times) are in progress.

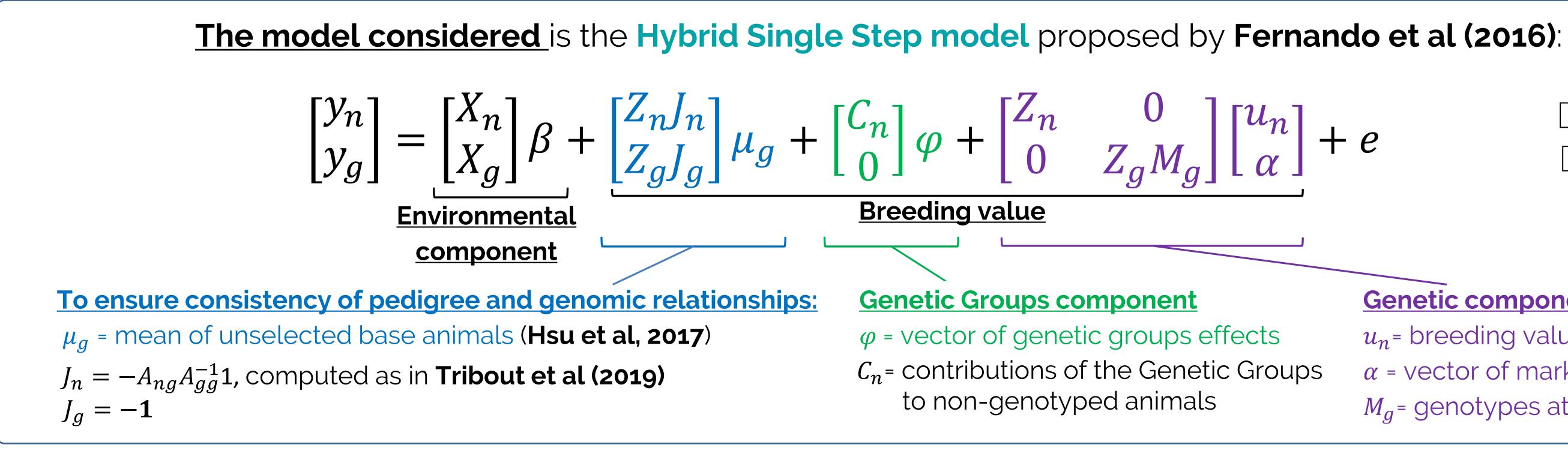
- All new evaluations have already been implemented with HSSGBLUP (e.g. see poster « Toward a genomic evaluation of cheesemaking traits including candidate SNP in Montbéliarde cows » #110 by Sanchez et al).
- All current French bovine polygenic and multistep genomic evaluations will be progressively replaced by Single Step SNP BLUP evaluations before april 2022 (dairy populations) and april 2023 (beef populations).

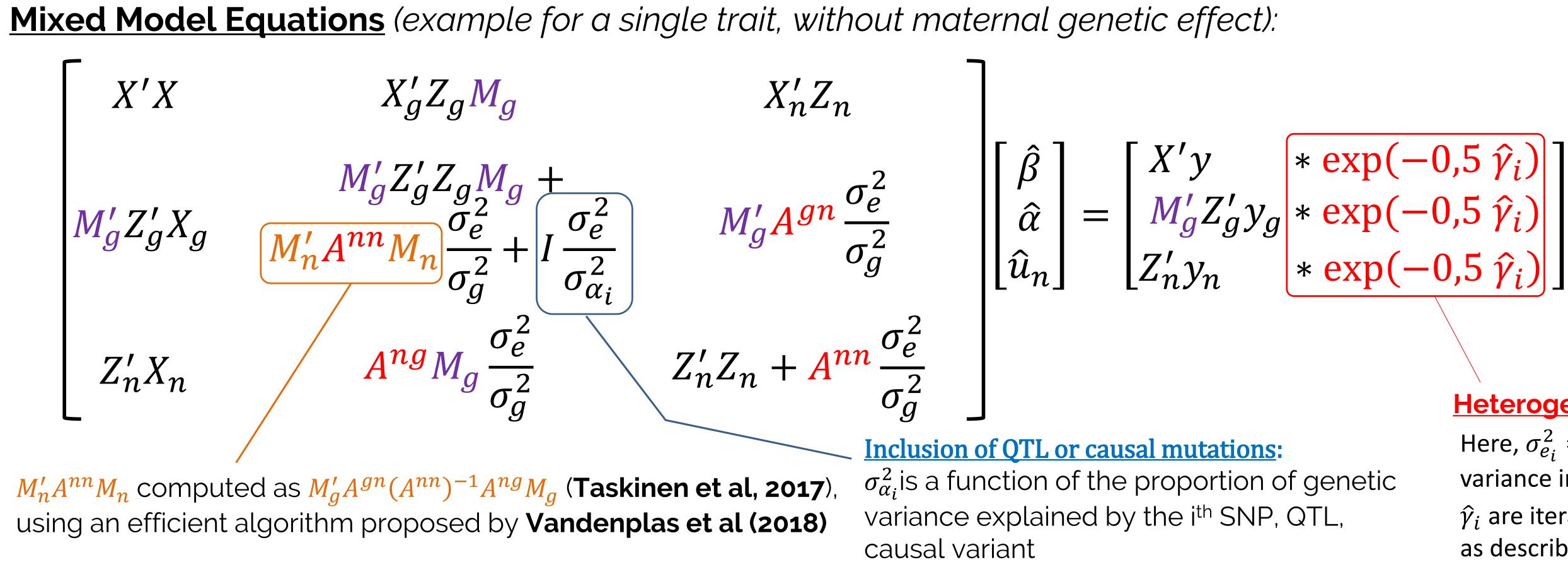
References

- Fernando L.R., Cheng H., Golden B.L. Garrick D.J., Genet. Sel.
- Hsu W. L., Garrick D.J., Fernando R.L., G3 (Bethesda) (2017) 7(8

> HSSGBLUP: a Single-Step SNP BLUP genomic evaluation software adapted to large livestock populations

Tribout T.¹, Ducrocq V.¹, Boichard D.¹





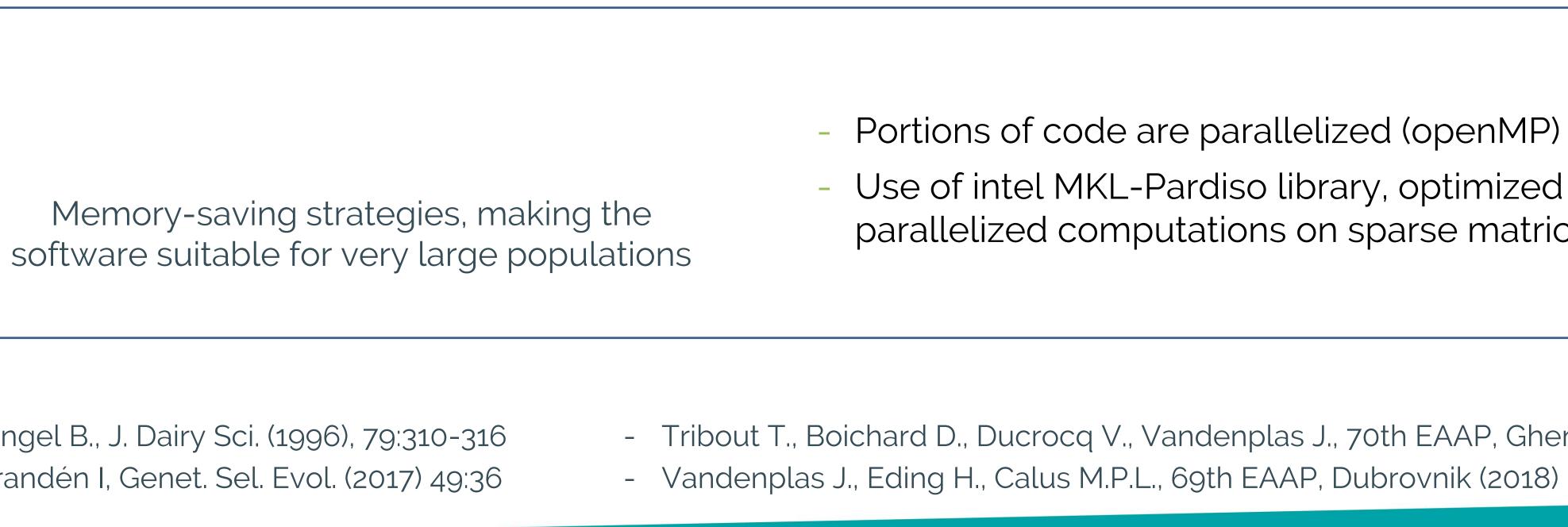
The genomic relationship matrix is neither built nor inverted \rightarrow the model is well adapted for populations with hundreds of thousands of genotyped animals

Programming strategies

- Coded in Fortran 90
- Solver: Preconditionned Conjugate Gradient
- Iteration on data
- Use of sparse matrices

Evol (2016) 48:96	_	Meuwissen T.H.E., De Jong G., Engel B., J. D
(8):2685-2694	-	Taskinen M., Mantysaari E.A., Strandén I, Ge

¹Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France



 $]_n = \text{non genotyped animals}$ $\Box_{q} = \text{genotyped animals}$

Genetic component

 u_n = breeding value of non-genotyped animals α = vector of markers effects M_a = genotypes at markers of genotyped animals

- Inverse of pedigree relationship matrix $A^{-1} = \begin{bmatrix} A^{nn} & A^{ng} \\ A^{gn} & A^{gg} \end{bmatrix}$
- M_n = (imputed) genotypes at markers of non-genotyped animals
- $\sigma_e^2 =$ residual variance
- σ_a^2 = genetic variance
- $\sigma_{\alpha_i}^2$ = genetic variance associated to the ith SNP, QTL, causal variant

Heterogeneous variances:

Here, $\sigma_{e_i}^2 = \exp(\gamma_i) \sigma_e^2$ is the residual variance in the ith level of heterogeneity $\hat{\gamma}_i$ are iteratively estimated on the data, as described in **Meuwissen et al (1996**)

- Portions of code are parallelized (openMP) - Use of intel MKL-Pardiso library, optimized for parallelized computations on sparse matrices

- Tribout T., Boichard D., Ducrocq V., Vandenplas J., 70th EAAP, Ghent (2019)

Centre de Recherche de Jouy-en-Josas 78350 Jouy-en-Josas thierry.tribout@inrae.fr