



Comparison of ssGBLUP and ssGTBLUP using Nordic Holstein TD data



M. Koivula¹, I. Strandén¹, G. P. Aamand² & E. A. Mäntysaari¹

¹ Natural Resources Institute Finland (Luke), FI-31600 Jokioinen, Finland

² NAV Nordic Cattle Genetic Evaluation, 8200 Aarhus N, Denmark

Summary:

GEBVs and genetic trends from the ssGBLUP and ssGTBLUP approaches were the same, excluding the small residual polygenic effect. ssGTBLUP offers a computationally reasonable approach for solving genomic breeding values using the single-step method. Eigendecomposition can be used to further reduce computations in ssGTBLUP.

Introduction

ssGBLUP is computationally demanding with large number of genotyped animals. Mäntysaari et al. (2017) proposed an exact approach named ssGTBLUP. The ssGTBLUP has the same MME as the original single step but the preconditioned conjugate gradient (PCG) iteration is made computationally less demanding by handling the \mathbf{G}^{-1} in factored form. The cost of ssGTBLUP scales linearly with respect to the number of genotyped animals, while the increase in ssGBLUP is quadratic.

The **aim of this study** was to compare the efficiency of the original ssGBLUP and ssGTBLUP approach with Nordic Holstein test-day (TD) data and model. The data had 7.6 million cows with 153 million records and 9.8 million animals in the pedigree. Genotype data included 101004 genotyped animals.

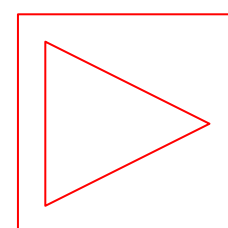
MiX99 ↩

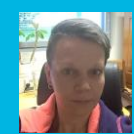
Studied single-step approaches

1. ssGBLUP_{w10}
"Traditional" single-step, 10% polygenic variance included into \mathbf{G} matrix top prevent singularity
2. ssGBLUP_{0.01}
 \mathbf{G} matrix formed similarly, but singularity constraint done by adding a diagonal matrix $\mathbf{C} = \mathbf{I}\epsilon$ where a small number ϵ was 0.01
3. ssGTBLUP
 \mathbf{G}^{-1} replaced by $\mathbf{C}^{-1}\mathbf{T}'\mathbf{T}$, where the \mathbf{T} matrix calculated for the ssGTBLUP approach using $\mathbf{C} = \mathbf{I}\epsilon$, and neither \mathbf{G} nor its inverse was explicitly formed
4. ssGTBLUP₍₉₈₎
Eigendecomposition used to reduce rank of the \mathbf{T} matrix. 98% of total variance in \mathbf{G} explained by the 14038 eigenvalues



Alex Arkink





Comparison of ssGBLUP and ssGTBLUP using Nordic Holstein TD data



M. Koivula¹, I. Strandén¹, G. P. Aamand² & E. A. Mäntysaari¹

¹ Natural Resources Institute Finland (Luke), FI-31600 Jokioinen, Finland

² NAV Nordic Cattle Genetic Evaluation, 8200 Aarhus N, Denmark

Summary:

GEBVs and genetic trends from the ssGBLUP and ssGTBLUP approaches were the same, excluding the small residual polygenic effect. ssGTBLUP offers a computationally reasonable approach for solving genomic breeding values using the single-step method. Eigendecomposition can be used to further reduce computations in ssGTBLUP.

Results

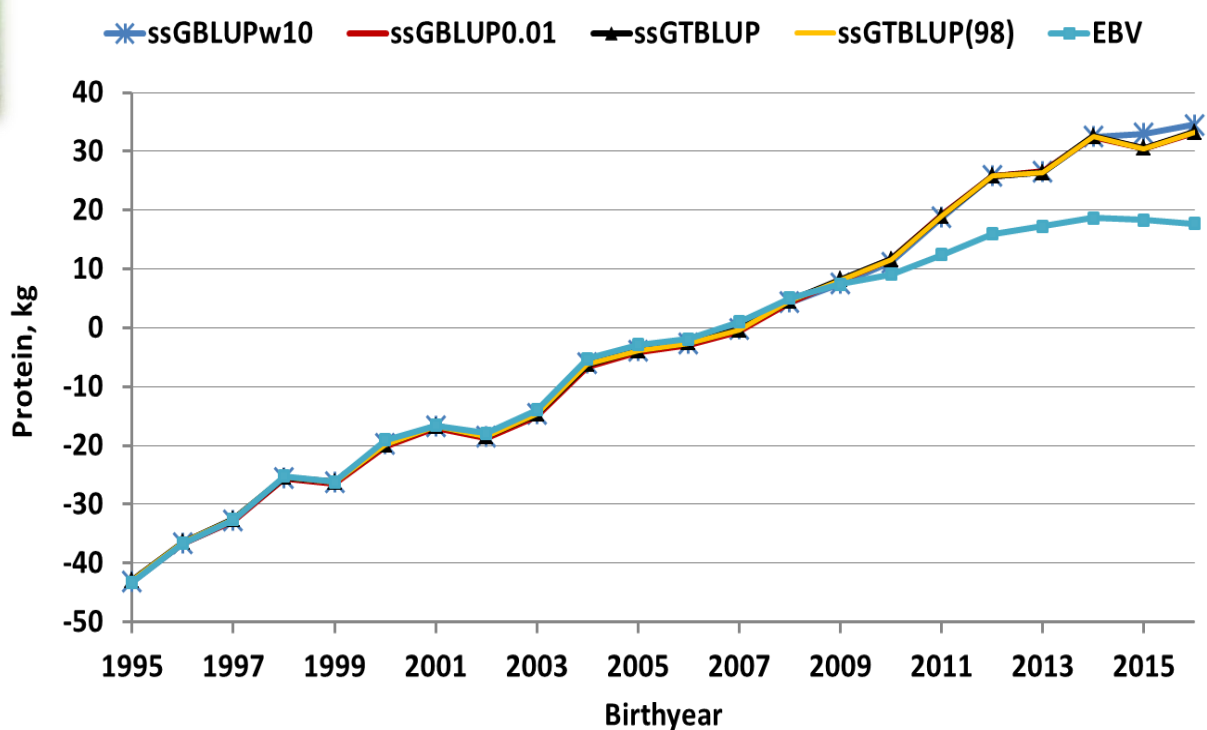
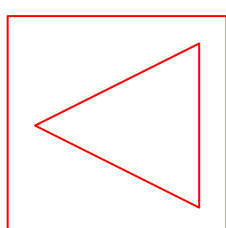
- ssGBLUP_{w10} needed the most pre-processing time and memory → A_{22} matrix had to be formed and inverted for the external genomic matrix
- For ssGTBLUP₍₉₈₎ eigendecomposition is more time-consuming than inverting a matrix of the same size
- Differences in ssGBLUP and ssGTBLUP methods become more important as the number of genotyped animals increases

Method	Peak memory (GB)	Pw (hour)	Matrix (GB)	I (h/1000 iterations)	N (Number)
Animal model	-	-	-	16	3290
ssGBLUP _{w10}	152.9	5.3	20	28	4202
ssGBLUP _{0.01}	114.7	2.2	20	25	4891
ssGTBLUP	51.8	1.5	18	25	4881
ssGTBLUP ₍₉₈₎	86.6	4	5.3	18	5576

Peak memory needed in gigabytes in the G/T matrix building. Wall clock time in hours for the preprocessing (Pw) and time per 1000 iterations in hours (I), size of external matrix read by the solver in gigabytes (Matrix), number of iterations (N).

Bull	Method	ssGBLUP _{0.01}	ssGTBLUP	ssGTBLUP ₍₉₈₎
Reference bulls	ssGBLUP _{w10}	1.00	1.00	1.00
	ssGBLUP _{0.01}		1.00	1.00
	ssGTBLUP			1.00
Young bulls	ssGBLUP _{w10}	0.98	0.98	0.98
	ssGBLUP _{0.01}		1.00	1.00
	ssGTBLUP			1.00

Correlations among GEBVs of genotyped Nordic Holstein bulls from single-step genomic models presented separately for reference bulls and for young bulls without daughters.



Genetic trends by birthyear for protein from different single-step approaches and TD animal model.