

Single-step genomic evaluation for fertility in Nordic Red dairy cattle

Kaarina Matilainen, Minna Koivula,
Ismo Strandén, Gert P. Aamand
and Esa A. Mäntysaari

Background

- Female fertility genetic evaluations (BLUP) have been done in Nordic countries since 1970's
- Joint Nordic fertility evaluations have been done since 2005
 - Nordic Cattle Genetic Evaluations NAV
- The model was upgraded in 2015
 - From sire to animal model
 - From repeatability to multi-trait model for lactations
- Next step: Genomic evaluation

Objectives of this presentation

- Single step genomic model (ssGBLUP):
Taking into account phenotypes, pedigree and genomic data simultaneously
- Genetic groups may cause problems in the convergence of the genomic model
 - QP-transformation for the full \mathbf{H}^{-1} matrix (unified relationships)
- Single-step genomic evaluation may need a long solving time
 - Algorithm for Proven and Young (APY)



Traits

- NAV fertility evaluations 2015 involve two different trait groups
- Model for *trait group* I contains 11 correlated traits
 - Heifer traits:
 - non-return rate (NRR0)
 - length of service period (IFL0)
 - Cow traits for lactations 1 - 3:
 - non-return rate (NRR1, NRR2, NRR3)
 - interval from calving to first breeding (ICF1, ICF2, ICF3)
 - length of service period (IFL1, IFL2, IFL3)

Data



- RDC data in routine joint Nordic fertility evaluations in 2016
 - Number of animals with observations: **4,226,636**
 - Number of animals in the pedigree: **5,445,392**
 - Number of genotyped animals: **33,969**
- Genetic parameters:
Estimated for the routine joint Nordic fertility evaluations in 2015
 - Low heritabilities (0.015-0.04)
 - High correlations among traits (0.60-0.88 between lactations)

Relationships in genomic evaluation

- In ssGBLUP the inverse of the relationship matrix is

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix}$$

where

- \mathbf{A} describes relationships based on the pedigree
- \mathbf{A}_{22} describes pedigree based relationships for genotyped animals
- \mathbf{G} gives relationships based on genomic information and

$$\mathbf{G}_w = (1 - w) \mathbf{G} + w \mathbf{A}_{22},$$

where w is the weight for polygenic information (we used 10%)

QP-transformation for genomic model

- Let rows in matrix \mathbf{Q} describe genetic group compositions for each animal
- Usually, \mathbf{A}^{-1} is augmented to include groups as phantom parents (PPG). This same transformation is blindly used in single-step:

$$\mathbf{H}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1} \\ -\mathbf{Q}'\mathbf{A}^{-1} & \mathbf{Q}'\mathbf{A}^{-1}\mathbf{Q} \end{bmatrix} & -\mathbf{A}^{-1}\mathbf{Q} \\ \mathbf{0} & \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} = \mathbf{A}_{ppg}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$

- However, contributions to PPG due to genomic relationships can be similarly accounted (Misztal et al., 2013):

$$\mathbf{H}^{-1} = \mathbf{A}_{ppg}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1} & -(\mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1})\mathbf{Q}_2 \\ \mathbf{0} & -\mathbf{Q}_2'(\mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1}) & \mathbf{Q}_2'(\mathbf{G}_w^{-1} - \mathbf{A}_{22}^{-1})\mathbf{Q}_2 \end{bmatrix}$$

where \mathbf{Q}_2 is a submatrix of \mathbf{Q} for the genotyped animals.

Algorithm for Proven and Young (APY)

- Increase in number of genotyped animals leads to computational challenges in forming, inverting and using the genotype based relationship matrix \mathbf{G} .

APY approach (Miszta et al., 2015):

- Divide \mathbf{G}_W to core (c) and young (y) animals
 - Approximate \mathbf{G}_W^{-1} inverse by
- $$\mathbf{G}_W = \begin{bmatrix} \mathbf{G}_{cc} & \mathbf{G}_{cy} \\ \mathbf{G}_{yc} & \mathbf{G}_{yy} \end{bmatrix}$$

$$\mathbf{G}_{APY}^{-1} = \begin{bmatrix} \mathbf{G}_{cc}^{-1} + \mathbf{G}_{cc}^{-1} \mathbf{G}_{cy} \mathbf{M}_{yy}^{-1} \mathbf{G}_{yc} \mathbf{G}_{cc}^{-1} & -\mathbf{G}_{cc}^{-1} \mathbf{G}_{cy} \mathbf{M}_{yy}^{-1} \\ -\mathbf{M}_{yy}^{-1} \mathbf{G}_{yc} \mathbf{G}_{cc}^{-1} & \mathbf{M}_{yy}^{-1} \end{bmatrix}$$

$$\mathbf{M}_{yy} = \text{diag}(\mathbf{G}_{yy} - \mathbf{G}_{yc} \mathbf{G}_{cc}^{-1} \mathbf{G}_{cy})$$

Here

- 12,741 animals that had descendant(s) were selected to the core
- Weight for polygenic information $w = 0.1$
- \mathbf{A}_{22}^{-1} is not formed explicitly

Analyses

Four genomic evaluations were performed

ssGBLUP	QP-transformation for pedigree based relationship matrix only
ssGBLUP _{QP}	QP-transformation both for pedigree and genomic information based relationships
ssGBLUP _{QP_Inb}	Like previous plus inbreeding coefficients taken into account in \mathbf{A}^{-1}
ssGBLUP _{QP_Inb_APY}	Like previous but QP- transformation for APY approximated genomic information based relationship matrix

Models were solved using MiX99,
and iterative preconditioned conjugate gradient algorithm (PCG)

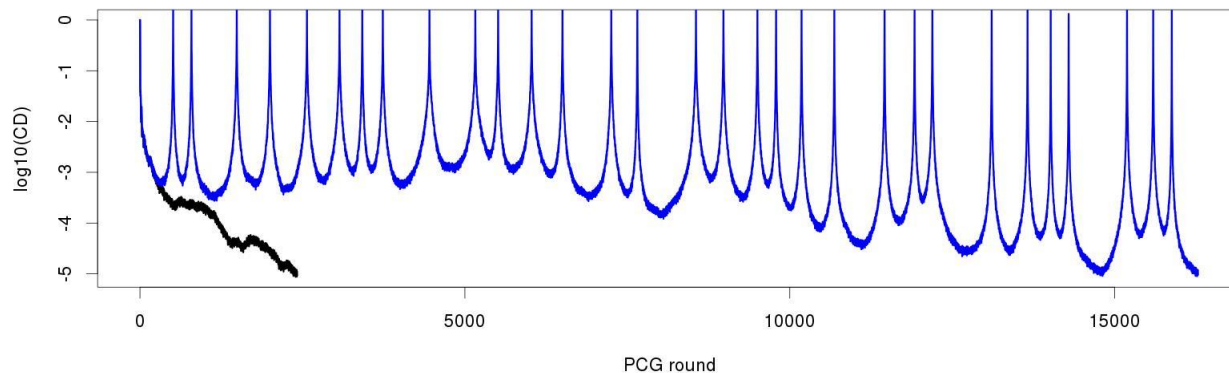
Comparisons between analyses

- Convergence
- Breeding values (for ICF2 and NRR3 shown as an example):
 1. Annual EBV and GEBV averages for males
 2. Annual EBV and GEBV correlations for both males and females
 3. Comparisons between GEBVs with and without APY
- GEBV validation tests



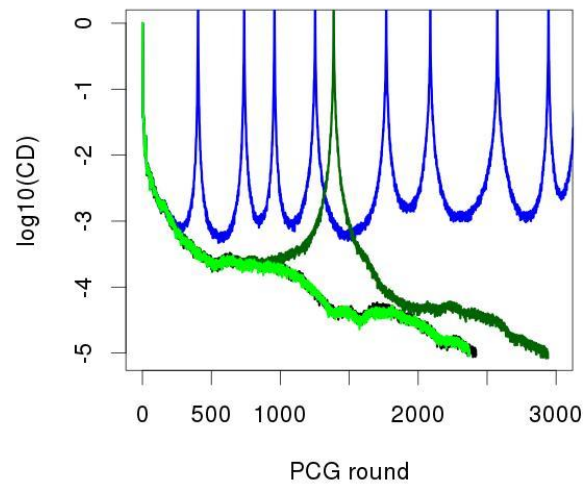
Convergence

Model	PCG rounds	Time	Time / round
BLUP	2,420	5h	7s
ssGBLUP	16,282	220h	49s



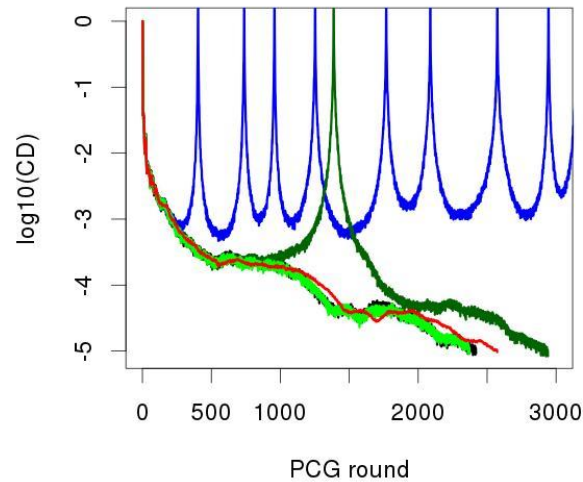
Convergence

Model	PCG rounds	Time	Time / round
BLUP	2,420	5h	7s
ssGBLUP	16,282	220h	49s
ssGBLUP _{QP}	2,941	45h	55s
ssGBLUP _{QP_Inb}	2,373	41h	62s



Convergence

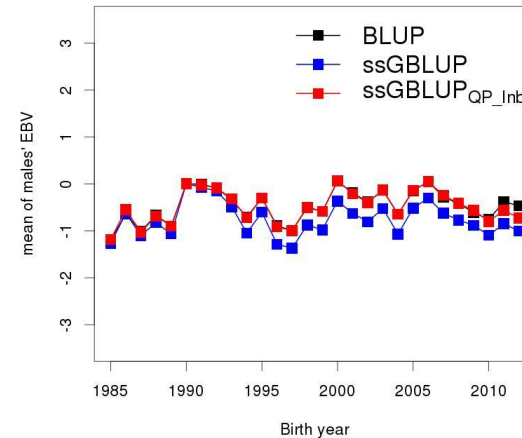
Model	PCG rounds	Time	Time / round
BLUP	2,420	5h	7s
ssGBLUP	16,282	220h	49s
ssGBLUP _{QP}	2,941	45h	55s
ssGBLUP _{QP_Inb}	2,373	41h	62s
ssGBLUP _{QP_Inb_APY}	2,573	34h	47s



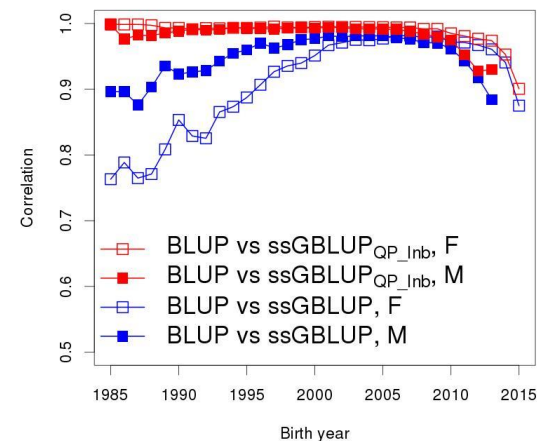
Comparison of GEBVs for ICF2 (Interval from calving to first breeding in second parity)

- After QP-transformation
 - Annual EBV and GEBV averages follows nicely each other.
 - Annual EBV and GEBV correlations were close to one for old animals but decreased somewhat for young animals.
- Correlations between GEBVs with and without APY were 1.000 and 0.998 for core and non-core animals, respectively.

Annual averages



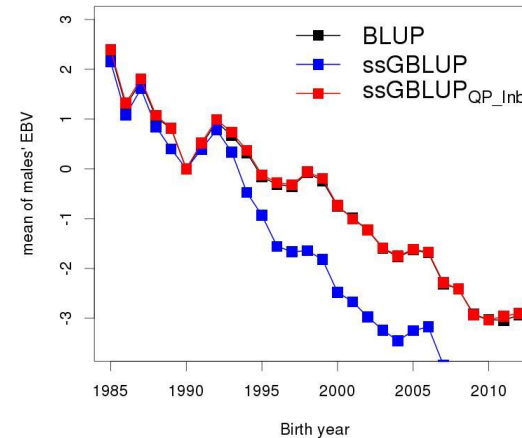
Annual correlations



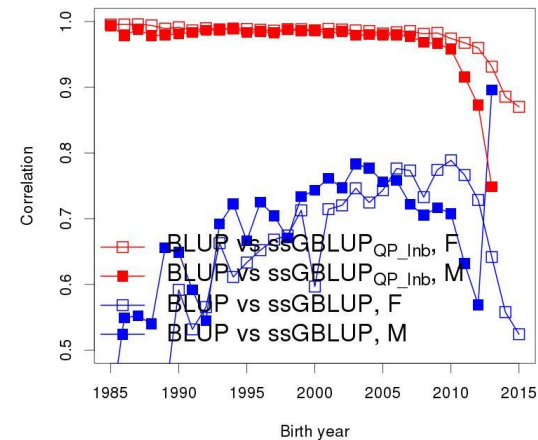
Comparison of GEBVs for NRR3 (non-return rate in third parity)

- After QP-transformation
 - Annual EBV and GEBV averages follows nicely each other.
 - Annual EBV and GEBV correlations were close to one for old animals but decreased somewhat for young animals.
- Correlations between GEBVs with and without APY were 1.000 and 0.999 for core and non-core animals, respectively.

Annual averages



Annual correlations



Interbull GEBV validation test

- Validation reliability (R^2) and regression coefficient (b_1) from the regression of deregressed genetic predictions from the full data on EBV and GEBV from the reduced data
 - **Observations from the latest 6 years were removed**
- Validation group contained 750 genotyped bulls
 - Bulls for which the effective record contribution: ERC > 10 based on full data and ERC = 0 based on reduced data



Validation reliabilities R^2 for ssGBLUP_{QP_Inb}

Trait	EBV	GEBV	$\Delta_{\text{GEBV-EBV}}$
NRR0	0.19	0.23	+0.04
IFL0	0.27	0.29	+0.02
NRR1	0.16	0.27	+0.11
ICF1	0.16	0.28	+0.12
IFL1	0.17	0.31	+0.14
NRR2	0.12	0.24	+0.12
ICF2	0.17	0.29	+0.12
IFL2	0.16	0.29	+0.13
NRR3	0.10	0.22	+0.12
ICF3	0.20	0.31	+0.11
IFL3	0.20	0.31	+0.11

NRR = Non-return rate
 IFL = Length of service period
 ICF = Interval from calving to first breeding

0 = Heifer
 1-3 = Parity

Regression coefficients b_1 for $ssGBLUP_{QP_Inb}$

Trait	EBV	GEBV	$\Delta_{GEBV-EBV}$
NRR0	1.00	0.81	-0.19
IFL0	1.06	0.87	-0.19
NRR1	0.96	0.86	-0.10
ICF1	0.99	0.90	-0.09
IFL1	0.92	0.89	-0.03
NRR2	0.98	0.95	-0.03
ICF2	0.88	0.86	-0.02
IFL2	0.85	0.89	+0.04
NRR3	0.83	0.92	+0.09
ICF3	0.92	0.90	-0.02
IFL3	0.88	0.91	+0.03

NRR = Non-return rate
 IFL = Length of service period
 ICF = Interval from calving to first breeding

0 = Heifer
 1-3 = Parity

Conclusions



- Single-step genomic evaluation for fertility in Nordic RDC was feasible.
- Accounting for genetic groups also in genomic information via QP-transformation was necessary:
 - Faster convergence
 - More consistent genomic breeding values when compared with traditional breeding values
- Considering inbreeding coef in \mathbf{A}^{-1} improved convergence greatly
- Model validation showed that ssGBLUP improved the fertility evaluations, especially for cow traits.
- APY-algorithm reduced the solving time with no effect on solutions

Acknowledgements

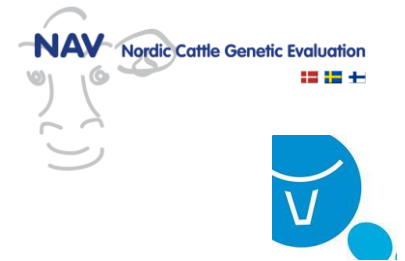
This work was part of the Nordic research project
GENOMICS in HERDS

Data:

NAV Nordic Cattle Genetic Evaluation

Funding:

MMM, NAV, Viking Genetics, Faba and Valio



Thank you!



NATURAL RESOURCES
INSTITUTE FINLAND