Improving single-step genomic prediction reliabilities for clinical mastitis in Nordic Red dairy cattle and Jersey by applying marker-specific weights

A. Chegini[1], I. Strandén[1], E. Karaman[2], T. Iso-Touru[1], J Pösö[3], G.P. Aamand[4], M. H. Lidauer[1]

[1]Natural Resources Institute Finland (Luke)
[2]Aarhus University, Denmark
[3]Faba co-op, Finland
[4]Nordic Cattle Genetic Evaluation, Denmark

Introduction

- In standard single-step genomic prediction equal SNP marker weights are assumed
- The true state may deviate from this assumption
- Markers may be in the range of influential genes or in proximity with them
- It has been suggested to consider SNP marker-specific weights in genomic prediction

Hypothesis

If SNP marker weighting is beneficial, it might be more pronounced in low heritable traits and traits having large QTLs

Objectives

- Applying SNP marker-specific weights in single-step genomic prediction and comparing it to standard single-step genomic prediction

- Testing with real data

  → Model and data of the Nordic udder health evaluation

Luke
NATURAL RESOURCES
INSTITUTE FINLAND

Data

- Breed-specific evaluations that include observations for clinical mastitis, test-day somatic cell score, and conformation traits
- Data from Denmark, Finland and Sweden that were recorded since 1980
- Nordic Red dairy cattle (RDC)

  75 million records from 5.6 million cows
- Jersey (JER)

  17 million records from 0.9 million cows

Trait definition

Clinical mastitis (CM) traits:

CM11: -15 to 50 days in milk, first lactation.

CM12: 51 to 305 days in milk, first lactation.

CM2: -15 to 150 days in milk, second lactation.

CM3: -15 to 150 days in milk, third lactation.

Correlated traits:

- 3 somatic cell score (SCS) traits: 1st , 2nd , 3rd lactation test-day observations (transformed to logarithmic scale).
- fore udder attachment (UA): 1st lactation.
- udder depth (UD): 1st lactation.

Luke
NATURAL RESOURCES INSTITUTE FINLAND

Summary statistics of studied traits

Trait	RDC			JER		
	n	Mean	SD	n	Mean	SD
CM11	4,791,842	0.06	0.25	601,988	0.14	0.35
CM12	4,641,064	0.06	0.24	590,198	0.11	0.31
CM2	3,452,089	0.11	0.31	427,327	0.14	0.34
CM3	2,246,733	0.14	0.35	287,409	0.16	0.37
SCS1	29,944,467	4.05	1.19	7,317,581	4.44	1.09
SCS2	20,997,798	4.41	1.28	5,100,038	4.67	1.19
SCS3	12,978,293	4.63	1.30	3,301,183	4.84	1.23
UA	1,161,944	5.58	1.36	307,635	5.47	1.17
UD	1,160,324	5.62	1.66	307,634	5.39	1.12

Genotype data

- Only genotypes from individual born after 2008 were used, leaving **64,777** and **125,789** genotyped JER and RDC, respectively
- 41,897 and 46,914 SNPs for JER and RDC, respectively

Genomic prediction model

- Multiple-trait random regression model (9 traits)
- Same effects as in the official evaluation model (Negussie et al., 2010)
- Additive genetic effects are modelled by covariance functions
- An animal's breeding values are modelled by 12 random regression coefficients which correspond to the 12 largest eigenvalues of the originally estimated additive genetic (co)variance matrix
- Single-step SNPBLUP for models with marker-specific weights
- Residual polygenic proportion: 10%

SNP marker weights

Own set of weights for each trait *j* (eigenvalue)

Weighting scenarios

1. **Nonlinear**:

$$w_{jk} = 1.25^{\frac{|\hat{u}_{jk}|}{\mathrm{sd}\,(\hat{u}_j)}-2},$$ (VanRaden, 2008; Cole et al., 2009)

2. $\mathbf{2pq\hat{u}^2}$:

$$w_{jk} = 2p_k q_k \hat{u}_{jk}{}^2,$$ (Falconer & Mackay, 1996)

3. **20SNP-window**:

Averaging the $\mathbf{2pq\hat{u}^2}$ weights of 20 adjacent SNP markers (Zhang et al., 2016)

Luke
NATURAL RESOURCES
INSTITUTE FINLAND

Model validation

- Validation method: Legarra & Reverter, 2018
- Reduced data: last four years of data were removed
- Validation group:
    - Bulls with ERC = 0.0 in reduced data and ERC ≥ 2.0 in full data
    - Cows with ERC = 0.0 in reduced data and ERC ≥ 0.9 in full data
- Validation traits: combined breeding values for CM and SCS
    - Weighting of traits:
        - CM combined (0.15, 0.15, 0.25, 0.45)
        - SCS combined (0.30, 0.25, 0.45)

Luke
NATURAL RESOURCES
INSTITUTE FINLAND

Validation results (Legarra & Reverter)

Clinical mastitis

Breed	Sex	N	Weight	b_0	b_1	R^2	gain%
RDC	**Male** (Female)	**86** (8440)	Standard ssGBLUP	**0.002** (0.005)	**0.75** (0.87)	**0.50** (0.74)	-
			Nonlinear	**0.001** (0.005)	**0.73** (0.85)	**0.51** (0.74)	**2.0** (1.1)
			$2pq\hat{u}^2$	**0.001** (0.003)	**0.68** (0.79)	**0.57** (0.78)	**13.8** (5.3)
			20SNP-window	**0.0004** (0.005)	**0.70** (0.85)	**0.49** (0.75)	**-1.6** (1.8)
JER	**Male** (Female)	**115** (8224)	Standard ssGBLUP	**0.013** (0.010)	**0.78** (0.89)	**0.65** (0.72)	-
			Nonlinear	**0.015** (0.012)	**0.77** (0.88)	**0.66** (0.73)	**0.5** (1.9)
			$2pq\hat{u}^2$	**0.010** (0.008)	**0.70** (0.79)	**0.66** (0.76)	**0.9** (5.3)
			20SNP-window	**0.012** (0.011)	**0.74** (0.87)	**0.64** (0.74)	**-2.4** (3.1)

Validation results (Legarra & Reverter)

Somatic Cell Score

Breed	Sex	N	Weight	b_0	b_1	R^2	gain%
RDC	Male (Female)	125 (18112)	Standard ssGBLUP	**3.40** (6.11)	**0.86** (0.97)	**0.58** (0.77)	-
			Nonlinear	**7.40** (6.84)	**0.83** (0.94)	**0.60** (0.78)	**2.6** (0.6)
			$2pq\hat{u}^2$	**7.21** (5.82)	**0.77** (0.87)	**0.64** (0.79)	**11.1** (2.3)
			20SNP-window	**6.66** (6.63)	**0.82** (0.94)	**0.59** (0.78)	**2.5** (1.0)
JER	Male (Female)	119 (6537)	Standard ssGBLUP	**8.17** (8.43)	**0.81** (0.97)	**0.61** (0.79)	-
			Nonlinear	**7.80** (8.43)	**0.80** (0.96)	**0.63** (0.80)	**2.7** (0.9)
			$2pq\hat{u}^2$	**4.06** (5.71)	**0.70** (0.87)	**0.65** (0.81)	**5.4** (2.8)
			20SNP-window	**7.55** (7.66)	**0.80** (0.95)	**0.64** (0.80)	**4.0** (1.3)

Conclusions

- SNP marker-specific weighting
    - improved bias and prediction reliability
    - but not dispersion
- Gain was higher in RDC than JER

BovReg *PARTNERS*

Thank you for your attention

www.bovregproject.eu

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 815668