Study of using genetic markers on a Nellore breeding program by model comparison

Fernanda Rezende; Ibáñez-Escriche, N.; Ferraz, J.B.S.;
Eler, J.P.; Silva, R.C.G.; Almeida, H.B.

62nd Annual Meeting EAAP, August 29th to September 2nd

Introduction

- Genetic evaluation
    - Pedigree + Phenotype, successful approach
    - DNA markers

- Brazilian beef cattle
    - 180 million animals, 80% *Bos indicus* contribution
    - Nellore breed
    - Associations genetic markers x quantitative traits

Objective

First evaluation of MAS implementation on a Nellore breeding program, applying different evaluation models and comparing those approaches in terms of model global fit and predictive ability for production traits.

Materials and methods

- **Phenotypic data**
  - Weaning weight (WW)
  - Post-weaning gain (PWG)
  - Scrotal circumference (SC)
  - Muscle score (MS)

  - Effects: contemporary group, age at measurement, age of dam at delivery, genetic group, management group at weaning and maternal
    - **VCE** (Groenevelt, 2008) { 138,565 animals measured; 235,766 animals at A^{-1}

Materials and methods

- **Descriptive statistics of analyzed dataset**

Trait	N	Mean	SD	CV
Weaning weight (WW, kg)	2,324	201.86	23.96	11.87
Post-weaning gain (PWG, kg)	2,313	120.67	32.33	26.79
Scrotal circumference (SC, cm)	1,894	27.18	3.14	11.56
Muscle score (MS, unit)	2,313	3.47	1.11	32.10

Materials and methods

- Genotypic data
  - ▫ 240 SNP markers chip
  - ▫ Filtering process: MAF > 5%
  - ▫ 103 analyzed markers

- Statistical models
  - ▫ Model 1: $y = \mu + Za + e$
  - ▫ Model 2: $y = \mu + Xg + e$
  - ▫ Model 3: $y = \mu + Xg + Za + e$

Materials and methods

- **Statistical analysis**
  - MCMC method (Gibbs algorithm)

$$y|\theta \sim N(\mu_j, I\sigma^2_e)$$

where the subscript j denotes the model (1, 2 or 3); θ are the unknown parameters; $\mu_1 = \mu + Za$; $\mu_2 = \mu + Xg$; $\mu_3 = \mu + Xg + Za$

- Prior distributions
  - $\mu \sim$ Uniform
  - $a \sim N(0, A\sigma^2_a)$
  - $e \sim N(0, I\sigma^2_e)$
  - $\sigma^2_a, \sigma^2_g, \sigma^2_e \sim$ Improper

Models 2 and 3:
$g \sim$ Uniform
$g \sim N(0, I\sigma^2_g)$

- TM program (Legarra et al., 2008), modified

Materials and methods

- **Rank correlation of model 1 with models 2 and 3**
  - Spearman correlation(ρ_s)
  - Kendall correlation (τ_b)

- **Model comparison**
  - Deviance Information Criteria (DIC) – global adjust
  - K-fold cross-validation – predictive ability
    - 4-fold cross-validation
    - 1-fold cross-validation

Mean squared error (MSE)
Pearson correlation (PC)

Results and discussion

- Posterior means and 95% highest posterior density intervals for additive variance of analyzed traits from models 1 and 3

Model	WW	PWG	SC	MS
1	38.96 (19.33, 59.08)	188.97 (128.90, 253.65)	3.21 (2.21, 4.29)	0.27 (0.16, 0.39)
3_U	33.58 (15.00, 54.48)	166.79 (107.33, 230.84)	2.89 (1.82, 3.90)	0.24 (0.12, 0.36)
3_N	36.12 (17.24, 56.83)	171.06 (109.69, 228.46)	3.02 (2.00, 4.02)	0.24 (0.13, 0.35)

- **Spearman (ρ_s) and Kendall (τ_b) rank correlations between model 1 and models 2 and 3**

Correlation	Trait	Model 2_U	Model 2_N	Model 3_U	Model 3_N
ρ_s	WW	0.31	0.31	0.76	0.98
	PWG	0.33	0.33	0.90	0.96
	SC	0.36	0.36	0.94	0.99
	MS	0.30	0.30	0.85	0.95
τ_b	WW	0.21	0.21	0.56	0.87
	PWG	0.23	0.22	0.73	0.82
	SC	0.25	0.24	0.78	0.90
	MS	0.20	0.20	0.65	0.81

Model comparison

- Estimates of deviance information criterion (DIC) from models 1, 2 and 3

Model	WW	PWG	SC	MS
1	2,710.98	3,147.96	2,648.33	2,902.87
2_U	2,446.83	2,437.42	2,007.19	2,417.38
2_N	2,372.71	2,394.40	1,950.02	2,380.95
3_U	2,751.40	3,169.14	2,660.99	2,941.82
3_N	2,705.68	3,156.77	2,647.82	2,928.52

- Mean estimates of mean squared error from models 1, 2 and 3, by 4-fold cross-validation

Trait	Model 1	Model 2_U	Model 2_N	Model 3_U	Model 3_N
WW	211.97	226.35	215.50	223.74	211.82
PWG	443.42	447.97	438.10	436.19	425.56
SC	6.74	7.40	7.01	7.07	6.67
MS	0.86	0.90	0.86	0.88	0.84

- Mean estimates of Pearson's correlation from models 1, 2 and 3, by 4-fold cross-validation

Trait	Model 1	Model 2_U	Model 2_N	Model 3_U	Model 3_N
WW	0.17	0.06	0.08	0.09	0.16
PWG	$\boxed{0.20}$	0.22	0.23	0.27	$\boxed{0.28}$
SC	0.25	0.14	0.16	0.23	0.27
MS	0.15	0.12	0.14	0.16	0.19

- Estimates of mean squared error from models 1, 2 and 3, by 1-fold cross-validation

Trait	Model 1	Model 2_U	Model 2_N	Model 3_U	Model 3_N
WW	140.48	152.99	141.38	154.93	141.83
PWG	481.13	506.49	487.53	505.62	486.33
SC	5.91	6.55	6.14	6.41	5.93
MS	1.25	1.38	1.32	1.34	1.29

- Estimates of Pearson's correlation from models 1, 2 and 3, by 1-fold cross-validation

Trait	Model 1	Model 2_U	Model 2_N	Model 3_U	Model 3_N
WW	0.05	-0.04	-0.06	-0.02	0.01
PWG	0.16	0.06	0.06	0.13	0.15
SC	0.19	0.07	0.08	0.14	0.18
MS	0.23	-0.02	-0.03	0.07	0.14

Summary

- **Model overall fit (DIC)**
  - Model 2

- **Model predictive ability**

4-fold cross-validation x 1-fold cross-validation

  - Random removed data – Model 3 (PWG)
  - Next generation performance – Model 1

Acknowledgements

Agro-pecuaria CFM Ltda.

Animal Breeding and Biotechnology Group

FAPESP

FUNPEC-RP

IRTA-Cataluña

University of Sao Paulo

Thanks for your attention

frezende@usp.br