
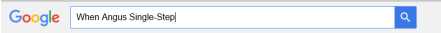
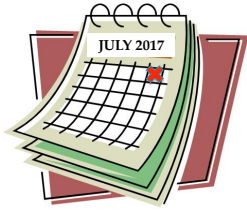


Genomic Prediction Developments at American Angus


Stephen Miller, Ph.D.
Kelli Retallick M.S.
Dan Moser, Ph.D.



The most asked Question

7/7



A roadmap for today

- Differences between "Production" and Single Step
- Practical implementation




Genomic Calibration

Genomic predictions in Angus cattle: Comparisons of sample size, response variables, and clustering methods for cross-validation¹


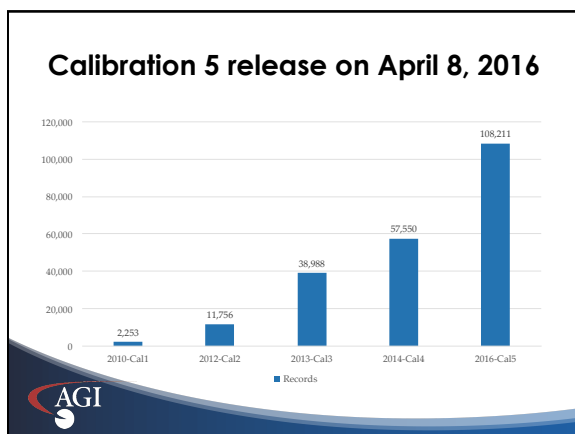
P. Boldorinsky,^{1,2} M. J. Kelly,³ S. Northercott,² K. C. Pryor,⁴ J. Rump,⁵ and S. DeNise⁶

¹Zoetis Inc., Kalamazoo, MI 49001; ²Quantitative Alliance for Agriculture and Food Innovation, University of Queensland, Brisbane St. Lucia, QLD 4072, Australia; ³American Angus Association, 1201 Franklin Ave., Suite 1000, 560-1400; and ⁴Zoetis Inc., 45 Poplar Road, Parkville, Victoria 3052, Australia

ABSTRACT: Advances in genomics, molecular biology, and statistical genetics have created a paradigm shift in the way livestock producers pursue genetic improvement in their herds. The series of these technologies has resulted in combining genotypic and phenotypic information to compute genomically enhanced measures of genetic merit of individual animals. However, large numbers of genotyped and phenotyped animals are required to produce robust estimates of the effects of SNPs that are assumed together to generate direct genomic breeding values (DGV). Data on 13,750 Angus animals genotyped with the Illumina BovineSNP50 Beadchip were used to develop genomic predictions for 17 traits reported by the American Angus Association through Angus Genetics Inc. in their National Cattle Evaluation program. Marker effects were computed using a 5-fold cross-validation approach and a Bayesian model averaging algorithm. The accuracies were examined with EBV and deregressed EBV (DEBV) response variables and with K-means and identical by state (IBS)-based cross-validation methodologies. The cross-validation accuracies obtained using EBV response variables were consistently greater than those obtained using DEBV (average correlations were 0.64 vs. 0.57). The accuracies obtained using K-means cross-validation were consistently smaller than accuracies obtained with the IBS-based cross-validation approach (average correlations were 0.54 vs. 0.64 with EBV used as a response variable). Comparing the results from the current study with the results from a similar study consisting of only 2,253 records indicated that larger training population size resulted in higher accuracies in validation animals and explained on average 18% (60% improvement) additional genetic variance across all traits.

Key words: accuracy, Angus cattle, clustering methods, cross-validation, identical by state, response variables

© 2014 American Society of Animal Science. All rights reserved. J. Anim. Sci. 2014.92:485–497. doi:10.2527/jas.2013-6757

Incorporation of Marker Scores into National Genetic Evaluations

Stephen D. Kachman^{*}
Department of Statistics, University of Nebraska-Lincoln

Abstract


As genetic tests for production traits become available it becomes important to incorporate them in our genetic evaluations. A model which treats marker scores as correlated traits is presented along with an approximate reduced model approach to make the model computationally feasible.

1 Introduction

Tests based on DNA marker panels have been coming online for a number of traits. The tests offer the potential to increase the reliability of genetic evaluation particularly for animals and traits with limited phenotypic information. To realize this goal it is imperative that information from these tests be incorporated into national genetic evaluations.

The number of markers in a marker panel can range from a single marker, to thousands of markers, to potentially the complete DNA sequence for an animal. Typically the results of a DNA test for a trait will be summarized into a single marker score or molecular breeding value. In most cases, the marker score will be a weighted sum of the number of copies of the different alleles with weights being estimated from a reference population. If the marker scores are on the same scale as the trait being evaluated, then the marker scores can be used directly in the genetic evaluation.

BIF 9th Genomic Prediction Workshop Kansas City MO, Dec 8-10, 2008



Implementing the Kachman model

	CWT_50K	
CWT	0.38	0.7
		0.99

Fit genomic prediction as a correlated trait



h2+g	RIB	CWT	LWT	URIB_B	URIB_C	URIB_S	CWT_M1	RIB_M1	CWT_M2	RIB_M2	CWT_50K	RIB_50K
RIB	X	X	X	X	X	X	X	X	X	X	X	X
CWT	X	X	X	X	X	X	X	X	X	X	X	X
LWT			X	X	X	X	X	X	X	X	X	X
URIB_B				X	X	X	X	X	X	X	X	X
URIB_C					X	X	X	X	X	X	X	X
URIB_S						X	X	X	X	X	X	X
CWT_M1							X	X	X	X	X	X
RIB_M1								X	X	X	X	X
CWT_M2									X	X	X	X
RIB_M2										X	X	X
CWT_50K											X	X
RIB_50K												X



Single Step allows a more DIRECT model

h2+g	RIB	CWT	LWT	URIB_B	URIB_C	URIB_S
RIB	X	X	X	X	X	X
CWT		X	X	X	X	X
LWT			X	X	X	X
URIB_B				X	X	X
URIB_C					X	X
URIB_S						X



Single Step allows a more COMPLETE model

h2+g	RIB	CWT	LWT	URIB_B	URIB_C	WWT	UFAT_B	UFAT_C	FAT
RIB	X	X	X	X	X	X	X	X	X
CWT		X	X	X	X	X	X	X	X
LWT			X	X	X	X	X	X	X
URIB_B				X	X	X	X	X	X
URIB_C					X	X	X	X	X
WWT						X	X	X	X
UFAT_B							X	X	X
UFAT_C								X	X
FAT									X



Is The Animal Model Obsolete?

L. R. Schaeffer

Centre for Genetic Improvement of Livestock
Department of Animal & Poultry Science
University of Guelph, Guelph, Ontario, Canada
Started: September 19, 2014
December 3, 2014

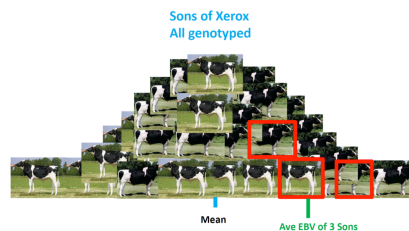
1 Introduction

The focus of animal breeding over the last 9 decades has been the evaluation of dairy bulls, and then cows. Thus, improving the models for genetic evaluation progressed from daughter-dam comparisons to herd-mate comparisons to contemporary comparisons to sire models and lastly to animal models. Much effort is spent on



Please tell me it isn't so !

A potential source of bias



Larry Schaeffer – University of Guelph 2014

University of Georgia

People

Professors

- [Luis Antonio Aguirre](#)
- [David L. Lusk](#)
- [David L. Lusk](#)

David L. Lusk

- [Luis Antonio Aguirre](#)
- [David L. Lusk](#)
- [David L. Lusk](#)

Specialist

- [Luis Antonio Aguirre](#)
- [David L. Lusk](#)
- [David L. Lusk](#)

Using recursion to compute the inverse of the genomic relationship matrix

A. Aguirre and L. Lusk

Journal of Dairy Science 94:2845-2852 (2011)

ABSTRACT

Computing the inverse of the genomic relationship matrix is a computationally intensive task. A recursive algorithm is proposed to compute the inverse of the genomic relationship matrix. The algorithm is based on the observation that the relationship matrix can be partitioned into blocks. The algorithm is implemented in the software package *BLUP*. The algorithm is implemented in the software package *BLUP*. The algorithm is implemented in the software package *BLUP*.

INTRODUCTION

Availability of genomic data has led to the development of genomic selection (GS). GS is a breeding strategy that uses genomic data to predict the breeding value of an individual. GS is a breeding strategy that uses genomic data to predict the breeding value of an individual. GS is a breeding strategy that uses genomic data to predict the breeding value of an individual.

Single Step GBLUP

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}W \\ W'R^{-1}X & W'R^{-1}W + H^{-1}\sigma_e^{-2} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ W'R^{-1}y \end{bmatrix}$$

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

W: incidence matrix of animals on data

A: pedigree relationship matrix

G: This G could be any matrix describing « genomic » covariances of breeding values; it does not restrict to VanRaden's (2008) GBLUP

A₂₂: pedigree matrix among genotyped individuals

Legarra, 2012

Single-Step Approach

- Redefines pedigree relationships among animals based on genetic markers

S

X

D

PGS

PGD

MGS

MGD

Aguilar et al, 2010

Genomics captures more variation in relationships

<p>6 - Full Sibs</p> <ul style="list-style-type: none"> • Pedigree – 0.59 • Genomics 0.49 – 0.65 	<p>232 – ½ sibs</p> <ul style="list-style-type: none"> • Pedigree min = 0.29 • Genomic min = 0.19
--	---

Thanks Daniela Laurencio

Single-Step incorporates all sources of information weekly

- ✓ Accounts for relationship between genotyped animals and genomic prediction information
- ✓ Eliminates periodic calibration
- ✓ Reduces bias

Pedigree & Genomic Data

↓

Performance

↓

Progeny Data

EPD

SS migration not unique to beef cattle in North America

- Software – UGA – Bolt – Mix99 – AGBU
- Beef – Dairy – Poultry – Pigs – Sheep

Current implementation size

- 7.8M birth weights
- 8.5M weaning weights
- 4.2M post-weaning gain records
- 1.5M heifer calving ease scores
- 113K carcass records
- 1.9M ultrasound records
- 20K individual intake records
- 254K docility scores
- 59K heifer pregnancy observations
- 202K mature cow weights
- 337K Genotypes



Re-estimated
h² estimates
and variance
components
increase
prediction
accuracy

Trait	Old h ²	New h ²
Calving Ease Direct	0.20	0.19
Birth Weight	0.42	0.43
Weaning Weight	0.20	0.20
Yearling Weight (Gain)	0.20	0.24
Dry Matter Intake	0.31	0.36
Yearling Height	0.50	0.51
Scrotal Circumference	0.47	0.48
Docility	0.37	0.44
Heifer Pregnancy	0.13	0.15
Milk	0.14	0.12
Mature Weight	0.37	0.37
Mature Height	0.64	0.62
Carcass Weight	0.38	0.44
Marbling Score	0.45	0.48
Ribeye Area	0.33	0.32
Fat Thickness	0.34	0.33

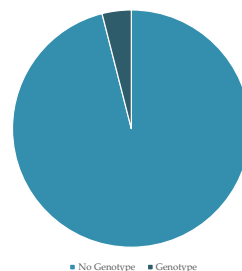


More Horsepower for Single Step

- 4 production servers running Linux with SSD hard drives
 - 1 – 24 CPU, 2.2ghz and 256gb memory
 - 3 – 12 CPU, 3.4ghz and 256gb memory
- 3X computing power
- 8X memory
- Expandable memory to grow with genotypes



Proportion of weaning weights genotyped



Changing horses on the fly



Running in Parallel – speed, correlation and stability



Photo credit FLICR Adrian Midgley

Comparison of EPDs

Trait	Top 200 sires	2,417 Young Sires
Calving Ease	0.92	0.85
Birth Weight	0.99	0.94
Weaning Weight	0.99	0.92
Yearling Weight	0.99	0.92
Milk	0.80	0.76
Mature	0.69	
Yearling Height	0.96	0.85
Scrotal	0.98	0.90
Heifer Pregnancy	0.67	0.60



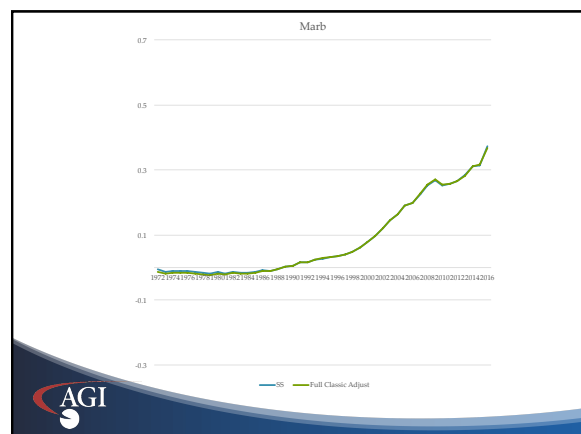
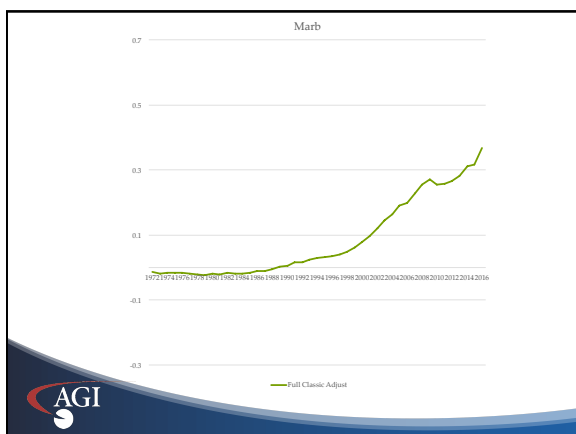
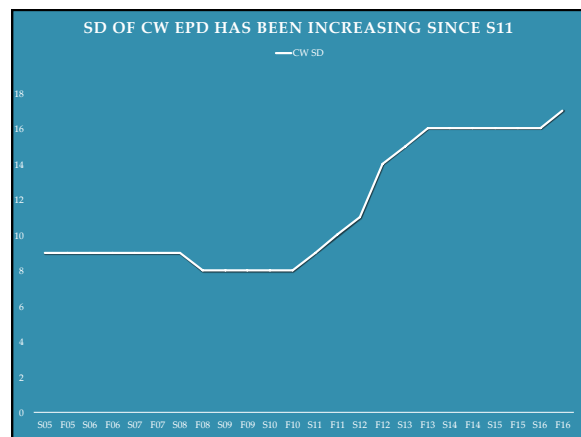
Current vs. Single Step

Comparison of EPDs

Trait	Top 200 sires	2,417 Young Sires
Carcass Weight	0.81	0.67
Marbling	0.86	0.85
Rib Eye	0.82	0.80
Fat	0.73	0.64

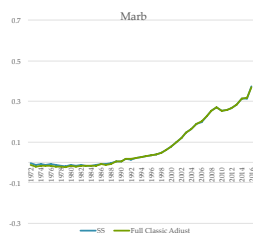


Current vs. Single Step



SS Genetic Trends more stable

- Inflating trends with MS
– contributing to drift
- SS and Classic trends very similar



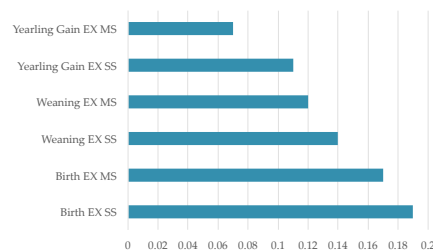
Single Step
proves
superior in
USMARC
validation

Correlations with MARC EPDs		
Trait	MS	SS
Birth	0.61	0.61
CWT	0.34	0.37
FAT	0.25	0.41
MILK	0.42	0.42
MARB	0.46	0.53
REA	0.25	0.23
Weaning	0.45	0.46
YWT	0.54	0.55

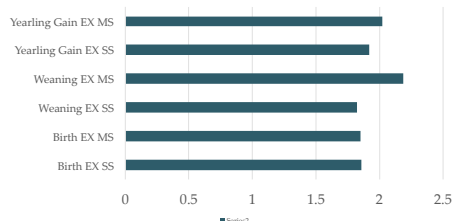
196 Sires



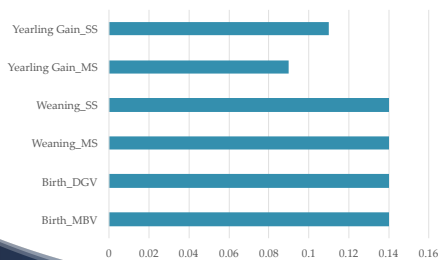
R-Square – SS is better



Regression coefficient – Close to expected (2)



R-Square DGV > MBV



Investigating YW Drift

- 12 bulls with YR ACC > 0.80 in 2011 still on the Top 200 list today.

	MS	SS
Regression lb/year	-2.1	-0.36
Change 2011- 2017	-13 lb	-2 lb

- Possible change for YW is 3.4 on a 0.8 Acc. Conclusion SS fixes YW drift.



EPD accuracy increases with Single Step – Weaning Weight

With Multistep, GE-EPD accuracy did not reflect relationships to other animals with data and genomics. Differences in accuracy are more readily seen in Single Step.

	Genotype		No Genotype	
Phenotype	Multistep:	0.32	Multistep:	0.25
	Single Step:	0.46	Single Step:	0.28
No Phenotype	Multistep:	0.28	Multistep:	0.02
	Single Step:	0.42	Single Step:	0.02



Progeny Equivalents

	Old H ²	New H ²	Old MS	New SS, FE	Ave. MS, Ave.	Ave. SS, Ave.
CE12	0.20	0.19	24	28	0.33	0.36
BW	0.42	0.46	13	21	0.36	0.48
WW	0.20	0.28	19	26	0.28	0.42
YW	0.20	0.27	24	21	0.32	0.38
Milk	0.14	0.12	14	33	0.19	0.30
DMI	0.31	0.34	17	8	0.34	0.24
HP	0.13	0.15	22	14	0.24	0.20
SC	0.47	0.48	16	13	0.39	0.40
YHT	0.50	0.51	12	11	0.36	0.39
Doc	0.37	0.44	11	10	0.28	0.34
MW	0.37	0.37	16	11	0.35	0.31
MH	0.64	0.62	7	8	0.18	0.36
CWT	0.38	0.44	7	9	0.31	0.34
MAEB	0.45	0.48	8	9	0.31	0.34
REA	0.33	0.32	11	12	0.27	0.30
Fat	0.34	0.33	12	10	0.30	0.29



Breeders can't wait for SS



Oversold?



Re-ranking and movement will not be “miniscule”

- 5 years in the making
- Benefits to breeders
 - Current
 - Complete
- Genotypes and Phenotypes required



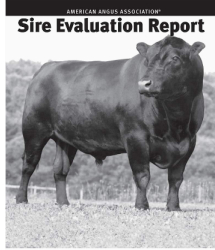
Structured Sire Evaluation is underway

- 1.8M Ultrasound - 110,000 Carcass records
- Carcass data from high use sires
- 600 cows bred in December 2015
- 1250 cows bred in Spring 2016
- 270 cows bred Fall 2016
- 3000 cows bred Spring 2017**
- Target large commercial cow-calf producers (250-600) with AI program



Several updates with the next *Sire Evaluation Report*

1. New methodology: *Single Step*
2. Re-estimated *variance components*
3. Carcass Rebuild
4. Incorporating *birth without weaning* data
5. Updated *Economic assumptions*



AGI



Questions

Dan W. Moser, Ph.D. dmoser@angus.org	AGI President 816-383-5196
Kelli Retallick, M.S. kretallick@angus.org	Genetic Service Director 816-383-5190
Stephen Miller, Ph.D. smiller@angus.org	Genetic Research Director 816-383-5157

AGI