ThetaSolutions

# Accuracies of contrasts between estimated breeding values of selection candidates from national cattle evaluations using pedigree or single-step genomic methodologies

D.P. Garrick[1,2], B.L. Golden[1], & D.J. Garrick[1,2]
1. Theta Solutions, LLC, USA
2. A.L. Rae Centre for Genetics & Breeding, Massey University
Daniel@ThetaSolutionsLLC.com

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

# EBV contrasts

- Evaluations produce individual EBVs for each trait (and typically a corresponding accuracy or reliability).

- Easy to find the "best" individual….
  - Rank ordering – e.g. sort by highest EBV (or index value).
    1. $EBV_{SireA}$ with accuracy/$R^2$.
    2. $EBV_{SireB}$ with accuracy/$R^2$.

- But how much better is it?
  - Contrasting EBVs of two (or more) animals.
    - $EBV_{SireA} - EBV_{SireB}$ with accuracy of?

# Single Trait Birth Wt. Model

- 2,118,874 animals in the pedigree.
- 1,416,006 birthweight observations.
- 38,175 genotypes using the MSRP subset (Saatchi & Garrick 2014).

- MMEs solved (PCG) and sampled (MCMC) using BOLT software.

- 80,000 MCMC samples of plausible values of every effect stored.

- EBVs are posterior means and PEVs are posterior variances of the chain of samples.

# Pedigree BLUP

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{u} + \mathbf{Z}_{\mathbf{m}}\mathbf{m} + \mathbf{Z}_{\mathbf{p}}\mathbf{p} + \mathbf{e}$$

- Fixed effects (**b**), direct effects (**u**), maternal genetic (**m**), maternal permanent environment (**p**), and residual (**e**) effects

# Super Hybrid Model

$$\begin{bmatrix} y_n \\ y_g \end{bmatrix} = \begin{bmatrix} X_n \\ X_g \end{bmatrix} b + \begin{bmatrix} Z_n & 0 \\ 0 & Z_g \end{bmatrix} \begin{bmatrix} M_n\alpha + \varepsilon \\ M_g\alpha \end{bmatrix} + Zu + Z_m m + Z_p p + e$$

- Super Hybrid Model (Fernando et al. 2016)
  for genotyped (**g**) and non-genotyped (**n**) animals.

- Includes marker effects (α).
- For non-genotyped animals, uses imputed markers ($\mathbf{M}_n$) and fits an imputation error term (**ε**).

# Accuracy of contrasts

- Calculated from diagonal and off-diagonal elements of the prediction error variance matrix.
  - PEV matrix is the inverse of the LHS of MME.
  - Inverse is computationally prohibitive, especially in single-step.

- Can approximate diagonal elements of inverse but PEV of contrasts rely on arbitrary off-diagonal elements

- Avoid approximation with MCMC (e.g. BOLT software).
  - GPU-accelerated single-site Gibbs sampler.

# Accuracy of Contrasts

- One column per animal of interest that contains its chain of plausible EBVs.
- Make chain of contrast of samples, e.g. $EBV_{SireA} - EBV_{SireB}$

$$\begin{bmatrix} EBV_{1,\ SireA} & EBV_{1,\ SireB} & \cdots & EBV_{1,\ SireZ} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ EBV_{N\ \ SireA} & EBV_{N\ \ SireB} & \cdots & EBV_{N\ \ SireZ} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \text{chain of } EBV_{SireA} - EBV_{SireB}$$

$\longrightarrow$ Contrast vector $\boldsymbol{k}$

- EBV of the contrast is the mean of the chain of contrasts.
- PEV of the contrast is the variance of the chain of contrasts.

# Results – Individual EBVs

- Results for “high” accuracy sires

High accuracy sires

| | PBLUP | | | SHM | | |
|---|---|---|---|---|---|---|
| ID | EBV | $R^2$ | BIF | EBV | $R^2$ | BIF |
| A | 0.71 | 0.97 | 0.84 | 0.84 | 0.97 | 0.84 |
| B | 3.04 | 0.96 | 0.81 | 3.15 | 0.97 | 0.81 |

# Results – Individual EBVs

- Genomic information improves low accuracy.

| | | PBLUP | | | SHM | | |
|---|---|---|---|---|---|---|---|
| | ID | EBV | $R^2$ | BIF | EBV | $R^2$ | BIF |
| High accuracy sires | A | 0.71 | 0.97 | 0.84 | 0.84 | 0.97 | 0.84 |
| | B | 3.04 | 0.96 | 0.81 | 3.15 | 0.97 | 0.81 |
| 2016 born males | C | 0.10 | 0.60 | 0.37 | -0.46 | 0.72 | 0.47 |
| | D | -0.03 | 0.59 | 0.36 | 0.36 | 0.72 | 0.47 |
| | E | 1.17 | 0.60 | 0.37 | 1.47 | 0.73 | 0.48 |
| | F | -1.86 | 0.61 | 0.37 | -1.81 | 0.73 | 0.48 |
| | G | 3.48 | 0.62 | 0.38 | [illegible] | [illegible] | 0.49 |
| | H | -0.03 | 0.59 | 0.36 | [illegible] | [illegible] | 0.47 |
| | I | -2.91 | 0.59 | 0.36 | [illegible] | [illegible] | 0.47 |
| | J | 0.95 | 0.59 | 0.36 | [illegible] | [illegible] | 0.47 |
| | K | -1.73 | 0.59 | 0.36 | -1.00 | 0.72 | 0.47 |
| | L | 1.17 | 0.60 | 0.37 | 1.47 | 0.73 | 0.48 |
| | M | -0.72 | 0.62 | 0.38 | -1.70 | 0.73 | 0.48 |
| | N | 0.95 | 0.59 | 0.36 | 1.55 | 0.72 | 0.47 |
| | O | 0.22 | 0.62 | 0.38 | 0.29 | 0.73 | 0.49 |
| | P | -1.21 | 0.58 | 0.35 | -0.88 | 0.70 | 0.45 |

Benefit of genotyping young animals

# Results – Contrasts

| Contrast | var(***k'u***) | PBLUP<br>PEV | <br>$R^2$ | <br>BIF | SHM<br>PEV | <br>$R^2$ | <br>BIF |
|---|---|---|---|---|---|---|---|
| B-A | 63.12 | 1.92 | 0.97 | 0.83 | 1.82 | 0.97 | 0.83 |

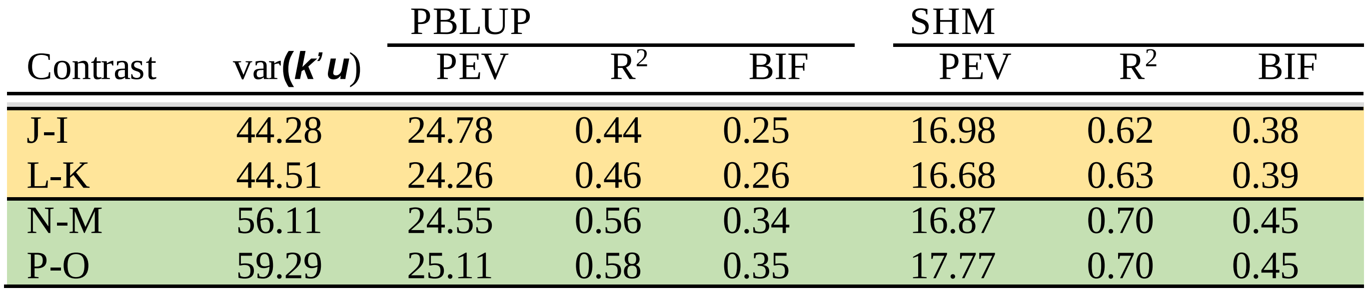
- High accuracy sires have high accuracy contrast (in this case)

# Results – Contrasts **same herd**

| | | PBLUP | | | SHM | | |
|---|---|---|---|---|---|---|---|
| Contrast | var($\boldsymbol{k'u}$) | PEV | $R^2$ | BIF | PEV | $R^2$ | BIF |
| D-C | 43.40 | 24.09 | 0.44 | 0.25 | 16.68 | 0.62 | 0.38 |
| C-E | 40.12 | 23.84 | 0.41 | 0.23 | 16.39 | 0.59 | 0.36 |
| G-F | 55.72 | 24.32 | 0.56 | 0.34 | 16.65 | 0.70 | 0.45 |
| H-G | 59.85 | 24.53 | 0.59 | 0.36 | 16.58 | 0.72 | 0.47 |

- **Young selection candidates with same sire**
- **Young selection candidates with different sires**

# Results – Contrasts **different herd**



| | | PBLUP | | | SHM | | |
|---|---|---|---|---|---|---|---|
| Contrast | var(***k*'*u***) | PEV | $R^2$ | BIF | PEV | $R^2$ | BIF |
| J-I | 44.28 | 24.78 | 0.44 | 0.25 | 16.98 | 0.62 | 0.38 |
| L-K | 44.51 | 24.26 | 0.46 | 0.26 | 16.68 | 0.63 | 0.39 |
| N-M | 56.11 | 24.55 | 0.56 | 0.34 | 16.87 | 0.70 | 0.45 |
| P-O | 59.29 | 25.11 | 0.58 | 0.35 | 17.77 | 0.70 | 0.45 |

- Young selection candidates with same sire
- Young selection candidates with different sires

- Genomic information improves accuracy of contrasts.
- Contrast between animals with the same sire have lower accuracy than those with different sires regardless of herd since PEV similar but var($\boldsymbol{k}'\boldsymbol{u}$) lower.
- These herds are well-connected due to wide AI use.

# Conclusions

- For comparing animals, it is the contrast (and the accuracy/reliability of contrasts) that matters.
- Accuracy of the individuals EBV’s are not an indication of the accuracy of the contrast.
  - Depends on prediction error co-variances which are influenced by “connectedness”.
- MCMC sampling of the MME using BOLT software is a computationally efficient method for national animal evaluations.
  - Same MCMC principles can be applied to selection indexes.

# Questions?

- Special thanks to the American Hereford Association for allowing the use of their national single-step evaluation data.

# EBV Accuracy

- Individual EBVs reported with a corresponding accuracy
  - Reliability $R^2$
  - Accuracy $r=\text{sqrt}(R^2)$
  - BIF accuracy

- “Measure” of the amount of information that went into producing the EBV
  - Quantify the possible variation of the EBV

- Prediction error variance (PEV)
  - Elements of inverse of LHS of MME
  - Inverse is prohibitive to compute, especially in single-step
  - PEVs commonly approximated

# Results - Contrasts

- For a contrast vector **k**, accuracy is computed as:
- $R^2 = 1 - \mathrm{PEV}_k / (\mathbf{k}'\mathbf{G}\mathbf{k})$ and BIF=1-sqrt(1- $R^2$)
- $\mathrm{Var}(\boldsymbol{k}'\boldsymbol{u}) = \boldsymbol{k}'\mathbf{G}\boldsymbol{k}$, where $\mathbf{G} = \mathrm{var}(\mathbf{u}) = \sigma_g^2 \mathbf{A}$ for single trait where **A** is the numerator relationship matrix

- For demonstration purposes $\mathrm{var}(\boldsymbol{k}'\boldsymbol{u})$ is taken to be the same in both the PBLUP and SHM (ie $\boldsymbol{k}'\mathbf{G}\boldsymbol{k}$)

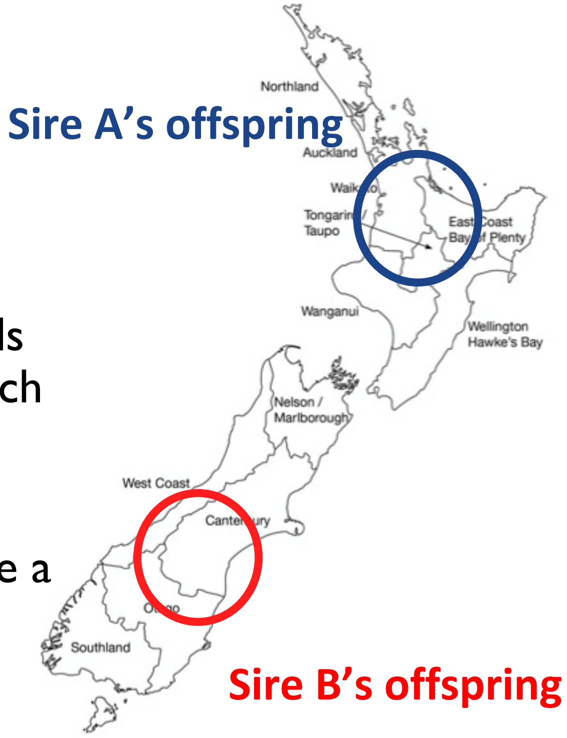
# Simplified hypothetical example

- Sire A and Sire B

- Both sires have many offspring.
- Both sires have high accuracy.

- But, offspring are in their own herds and geographically isolated from each other.

- A contrast of these sires might have a low or high accuracy.
  - Even in the absence of GxE



https://geog397.wiki.otago.ac.nz/images/9/92/MAP.jpg

# Overview

- Why EBV contrasts are important.
- Single trait birthweight model
  - Pedigree BLUP genetic evaluation.
  - Single-step Super Hybrid Model (SHM) genomic evaluation.
- Accuracy of contrasts.
- Results.
  - Contrast of high accuracy sires.
  - Contrasts of 2016 born males.
- Conclusions.