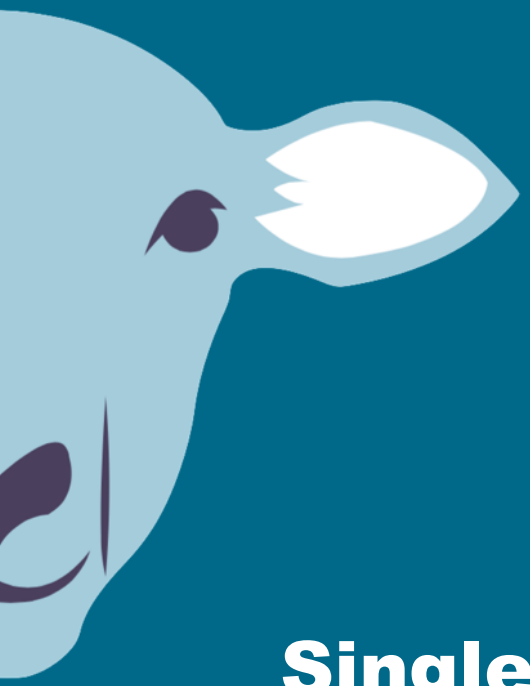


Single-Step genomic analyses:

Value to breeders and farmers

Dorian Garrick D.Garrick@massey.ac.nz
AL Rae Centre for Genetics and Biology
Ruakura Agricultural Research Centre

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

Underlying Concept

- Performance = Breeding + Feeding
- Phenotype = Genotype + Environment

Environment

- Phenotype = flock-year + age-of-dam + sex + BRR + BV + residual

- BV = Mα + EPE

Explained by markers

Not explained by markers (Extra Polygenic Effect)

What proportion of BV can be explained by markers?

- Eventually – some scientists believe it may be 100%
- Currently, with huge datasets, it is hard to exceed about 70%
  - This means that predictions of widely-used sires based on markers alone will not be the same as their progeny test predictions
  - This means that EBV based only on MBVs will not be as accurate as can be achieved with existing data
  - Accordingly, we want to “blend” the useful information we get from markers with the usual information we get from phenotypes on the individual or its close relatives
    - Computing the correct “weights” for the various sources of information is not trivial

Two-Step Approach

- Compute the marker effects from some historical data
- Use those marker effects on newly genotyped individuals (MBV)
  - Extend the MBV to non-genotyped relatives
- Compute the information from pedigree relatives (EBV)
- Combine the EBV and MBV into a GEBV by weighting each source
  - Using a Selection Index
  - By including the MBV as a “correlated trait” in the usual pedigree analysis

Selection Index Assumptions

beef+lamb GENETICS | SHEEP BREEDER FORUM

$$var\begin{bmatrix} u-\widehat{u} \\ m-\widehat{m} \end{bmatrix} = \begin{bmatrix} 1-r_p^2 & (1-r_p^2)(1-r_m^2) \\ (1-r_p^2)(1-r_m^2) & 1-r_m^2 \end{bmatrix}$$

$$var\begin{bmatrix} \widehat{u} \\ \widehat{m} \\ u \end{bmatrix} = \begin{bmatrix} r_p^2 & r_p^2 r_m^2 & r_p^2 \\ r_p^2 r_m^2 & r_m^2 & r_m^2 \\ r_p^2 & r_m^2 & 1 \end{bmatrix} \sigma_g^2$$

Blending

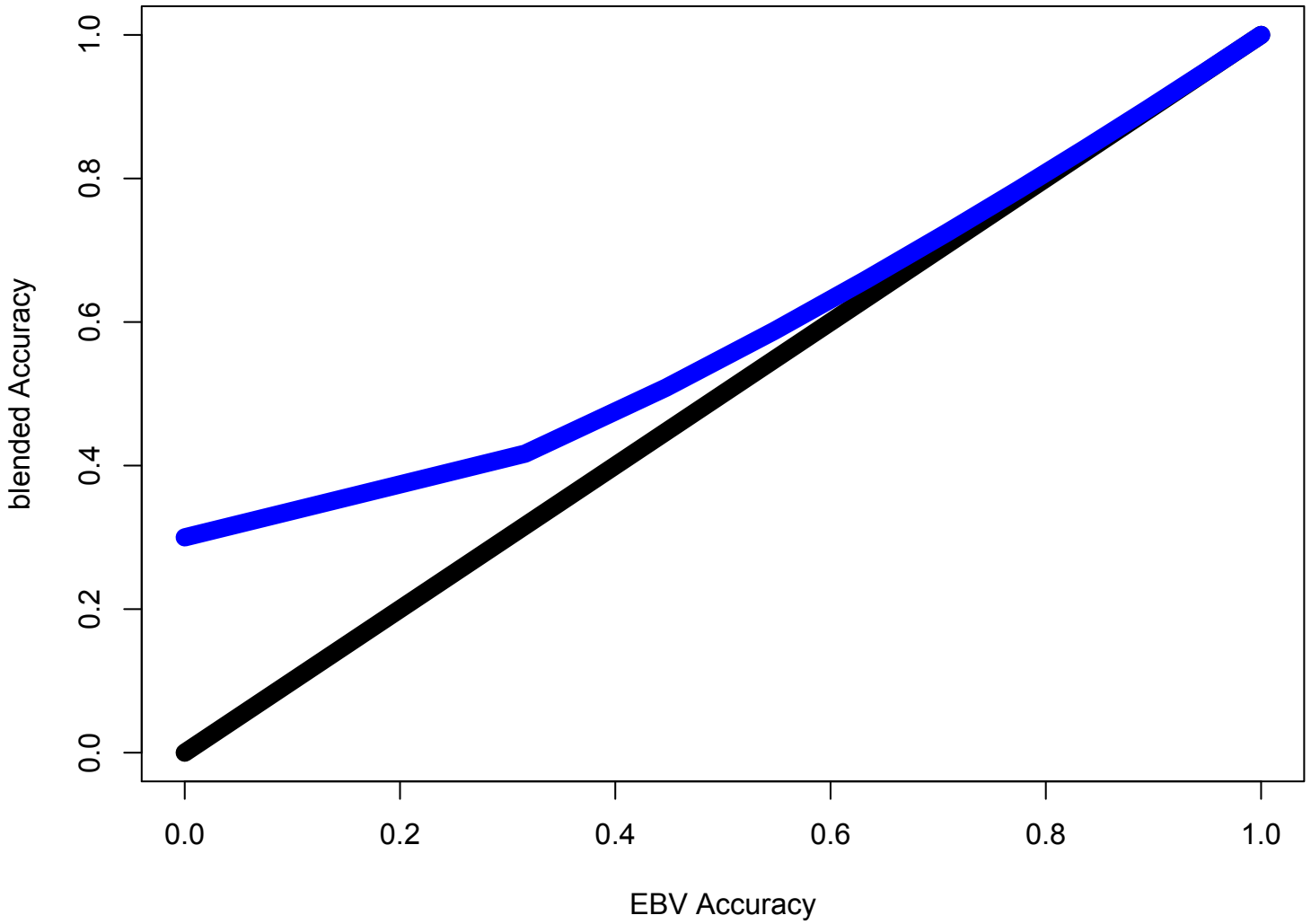
$$\widehat{u_n} = \frac{(1-r^2)(\widehat{u_p} - \mu_{u_p}) + (1-a^2)(\widehat{m} - \mu_m)}{1-r^2a^2}$$

$$Rel_n = 1 - \frac{(1-r^2)(1-a^2)}{1-r^2a^2}$$

where $\widehat{u_p}$ *is the previous national EBV with* $Rel_p = a^2$
and $\widehat{m}$ *is the MBV (DGV) with genetic correlation* r^2

Impact on Accuracy--%GV=10%

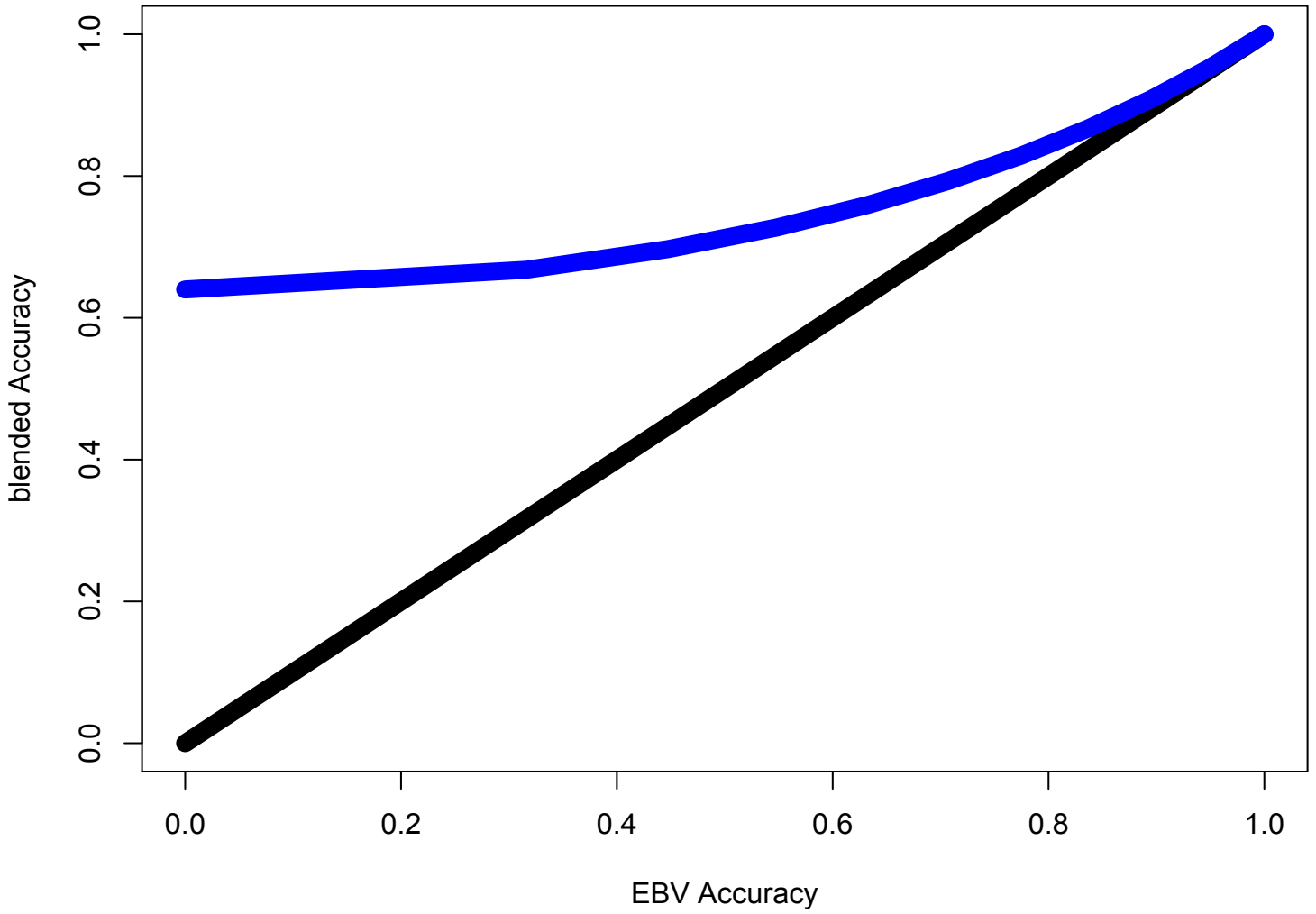
Genetic correlation=0.3



Blending will not improve the accuracy of an animal that already has a reliable EBV

Impact on Accuracy--%GV=40%

Genetic correlation=0.64



Blended EBVs are equally likely to be better or worse than the pre-blended EBV

beef+lamb GENETICS | SHEEP BREEDER FORUM

Problems

- The accuracy of all MBV are not equal – and are hard to calculate
  - Therefore the weights are not quite right
- If any records are available on newly genotyped animals these were not used in the computational of marker effects, so MBV will not be as accurate as possible
- GEBV may “jump” when marker effects are recomputed
  - As happens from time-to-time

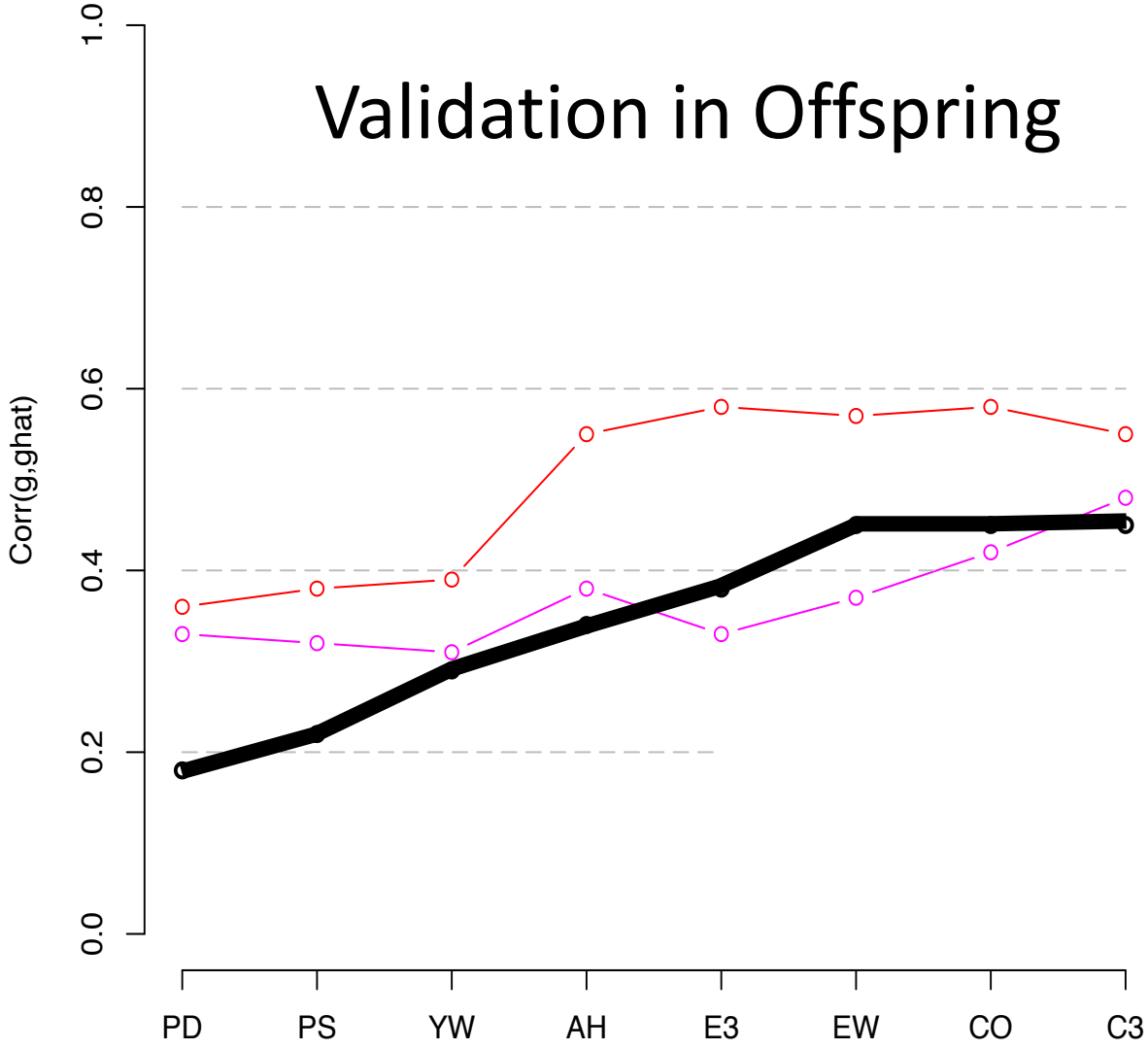
Solution is known as “Single Step”

- Fit a joint model (known as Single Step) that explicitly or implicitly includes the two components of the GEBV
  - Namely the part explained by markers and the part unexplained by markers
- There are many different ways to do “Single Step”
  - Single Step GBLUP (ssGBLUP) – use markers to form relationships
    - Works very well within-breed for small numbers genotyped
    - Requires approximations when >100,000 animals genotyped
    - More problematic to fit when data represents admixed breeds
  - Single Step Bayesian Regression (ssBR) – explicitly solve marker effects
    - Gets easier as a higher proportion of animals are genotyped
    - Can assume some markers have 0 effects (variable selection or mixture model)
    - Allows different markers for different traits
    - Different formulations for “most animals genotyped” or “most animals not genotyped”

Two-Step vs Single-Step

- Really a misnomer – all evaluations have multiple steps!
  - Generating a clean pedigree
  - Generating a clean file of performance data with cohort definitions
  - Generating a clean file of “imputed” quality-controlled genotypes
  - Perhaps pre-adjusting some of the phenotypes
  - Calculating the EBVs (Single-step vs Two-step – but also multiple traits)
  - Calculating the accuracies (most commonly from approximations)
  - Error-checking the results
  - Forming the index values
  - Distributing the results of the analyses

Accuracy of Genomic Prediction – Layer Hens

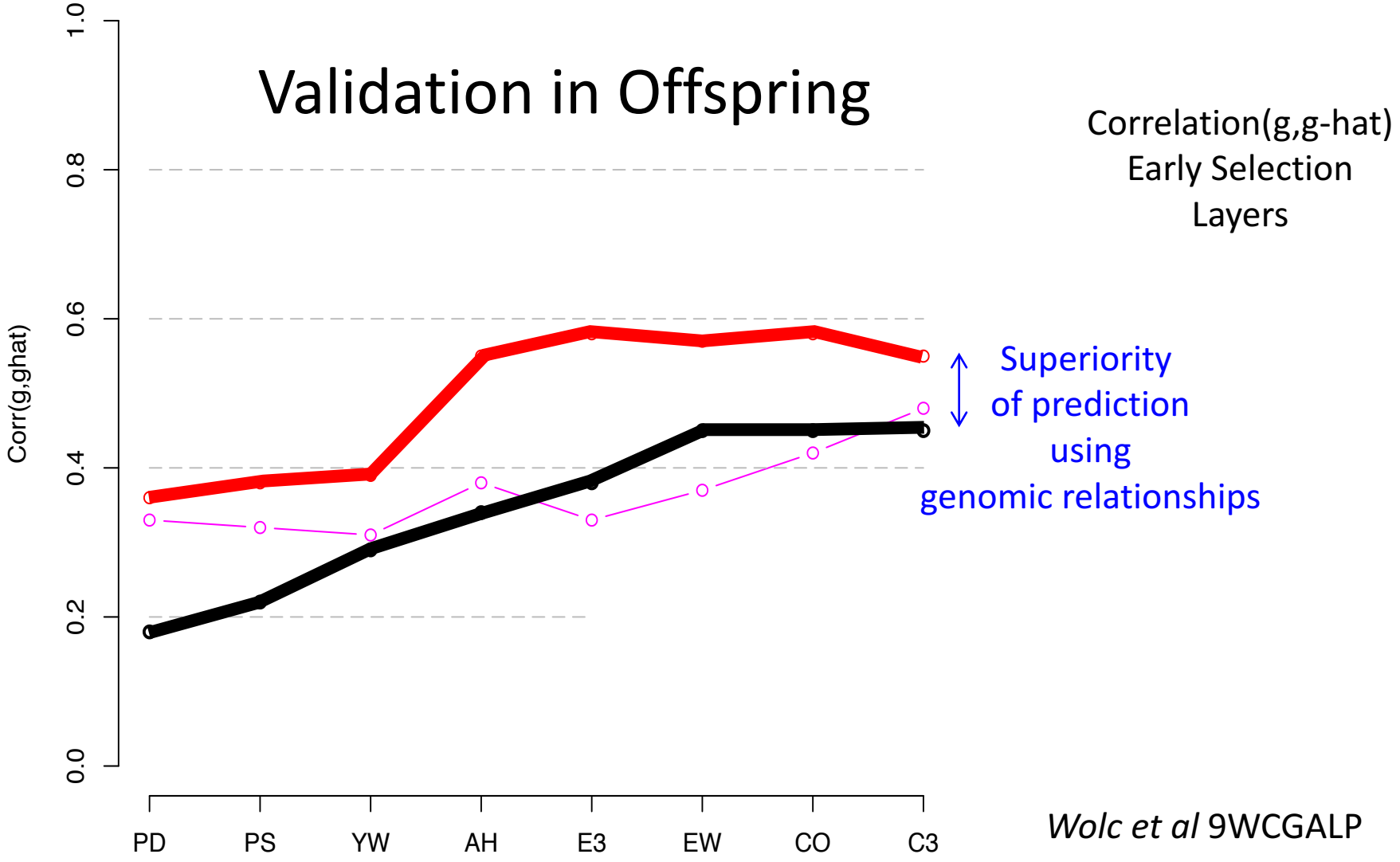


Correlation(g,g-hat)
Early Selection
Layers

Conventional
pedigree
relationships

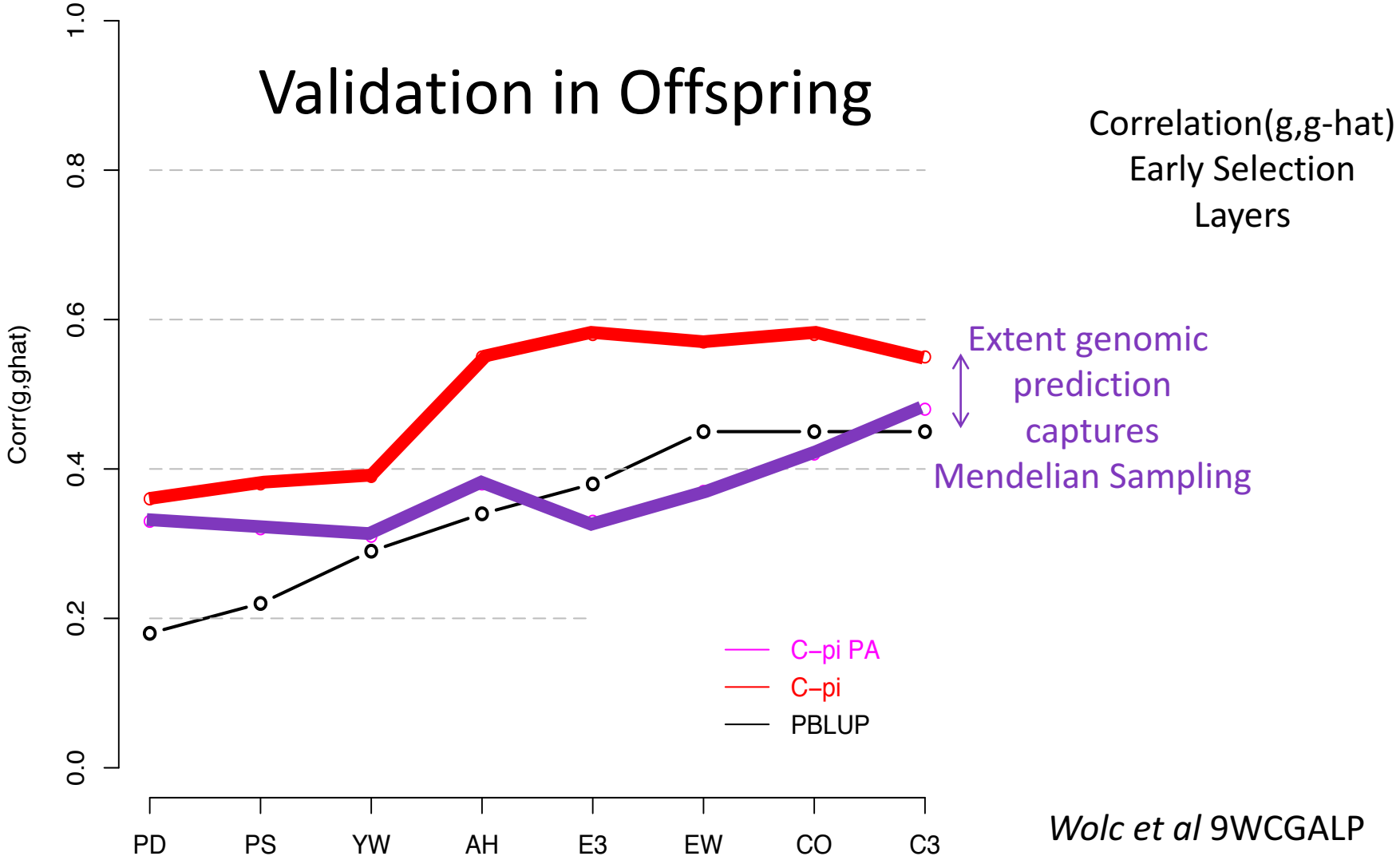
Wolc et al 2010 9WCGALP

Accuracy of Genomic Prediction – Layer Hens



Wolc et al 9WCGALP

Accuracy of Genomic Prediction – Layer Hens



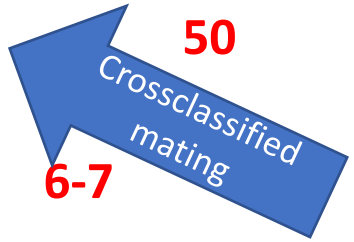
Wolc et al 9WCGALP

Layer Hens – Dekkers scheme

Strategy	Traditional	
	Male	Female
#candidates with phenotype	1000	3000
# selected	60	360
Generation interval (months)	13	
Information	Own Phenotype	

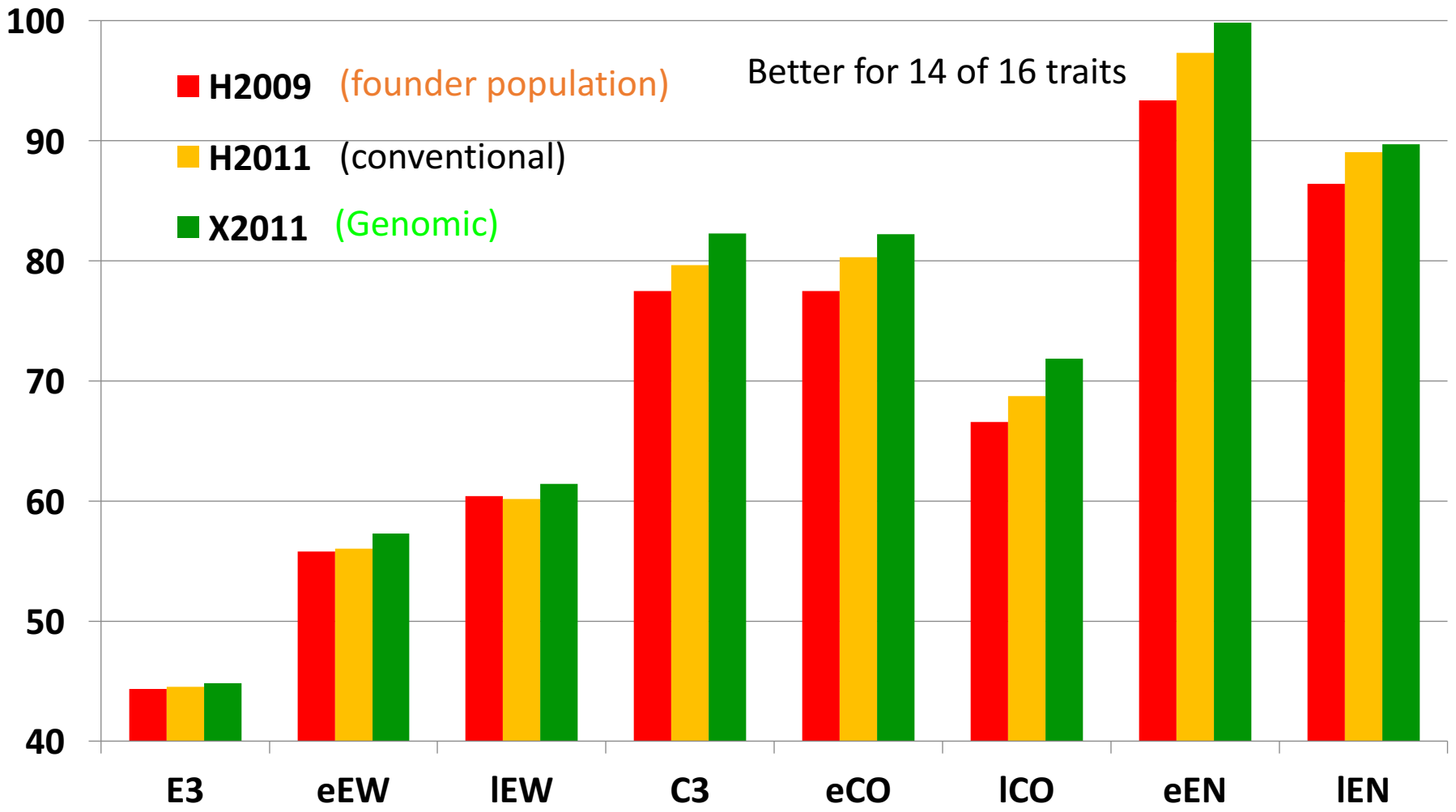
Layer Hens – Dekkers scheme

Strategy	Traditional		GS	
	Male	Female	Male	Female
#candidates with phenotype	1000	3000	300	300
# selected	60	360	50	50
Generation interval (months)	13		6-7	
Information	Own Phenotype		Genotype+Phenotype	



Could halve the generation interval and reduce costs by (less phenotyping) to get same gain & same inbreeding

Selection Response - Difference between the lines

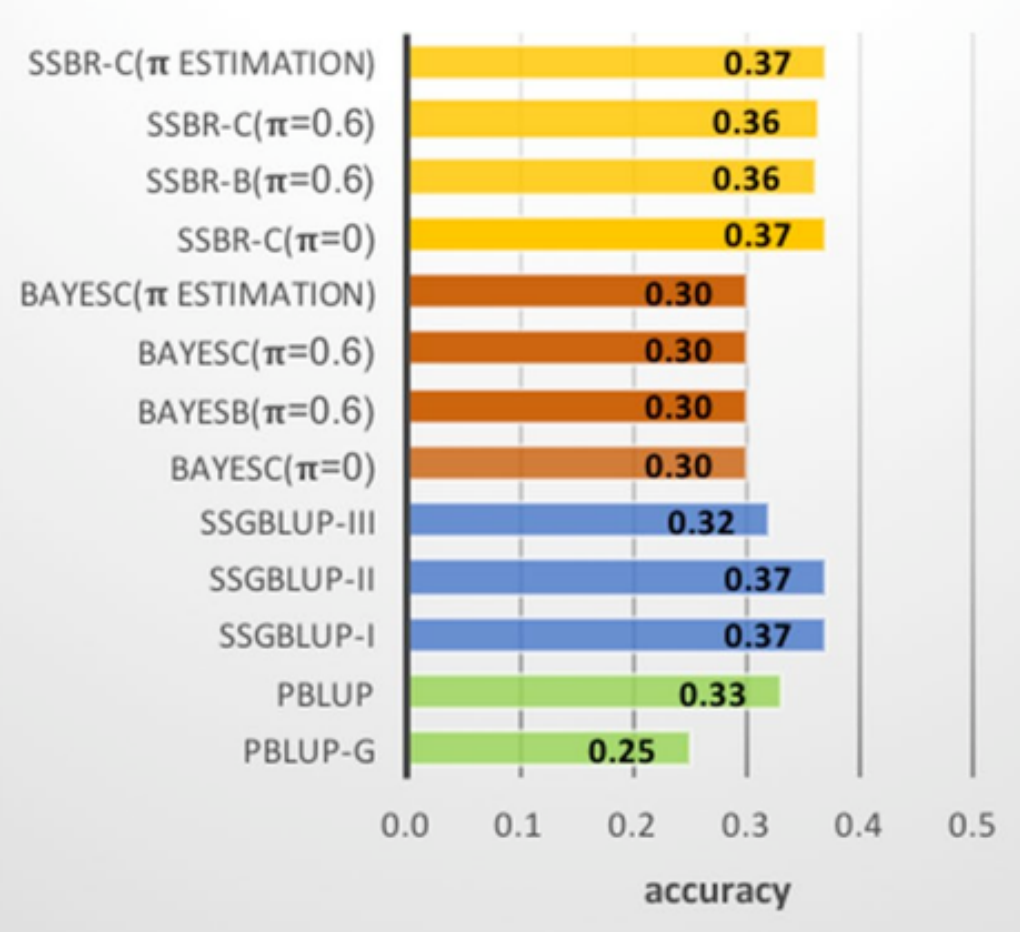


After 3 generations of conventional or 6 gens of genomic selection

Genomic selection was as good, if not better in terms of realized response

Wolc *et al.* Genetics Selection Evolution (2015) 47:59

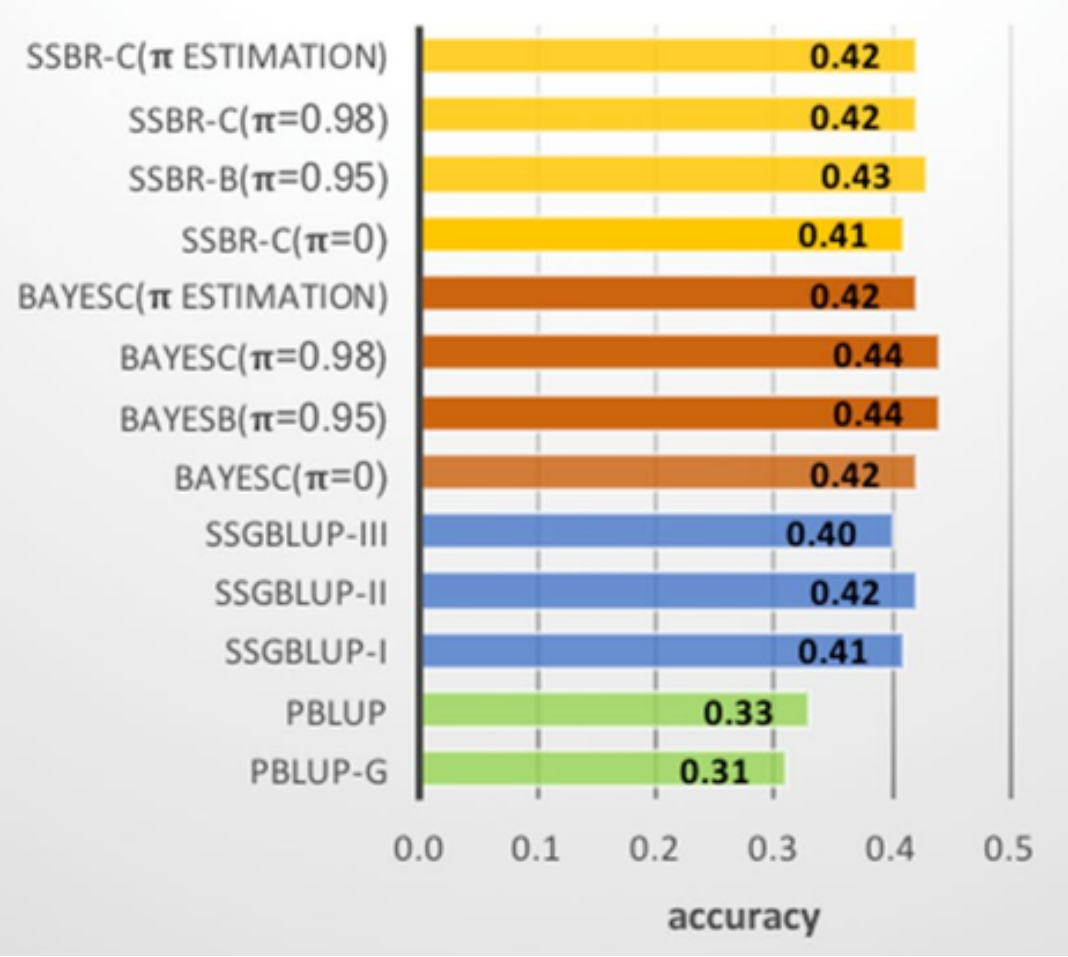
Hanwoo Marbling



All Single-Step methods outperformed pedigree methods and those only using genotyped

Lee *et al. Genet Sel Evol (2017) 49:2*

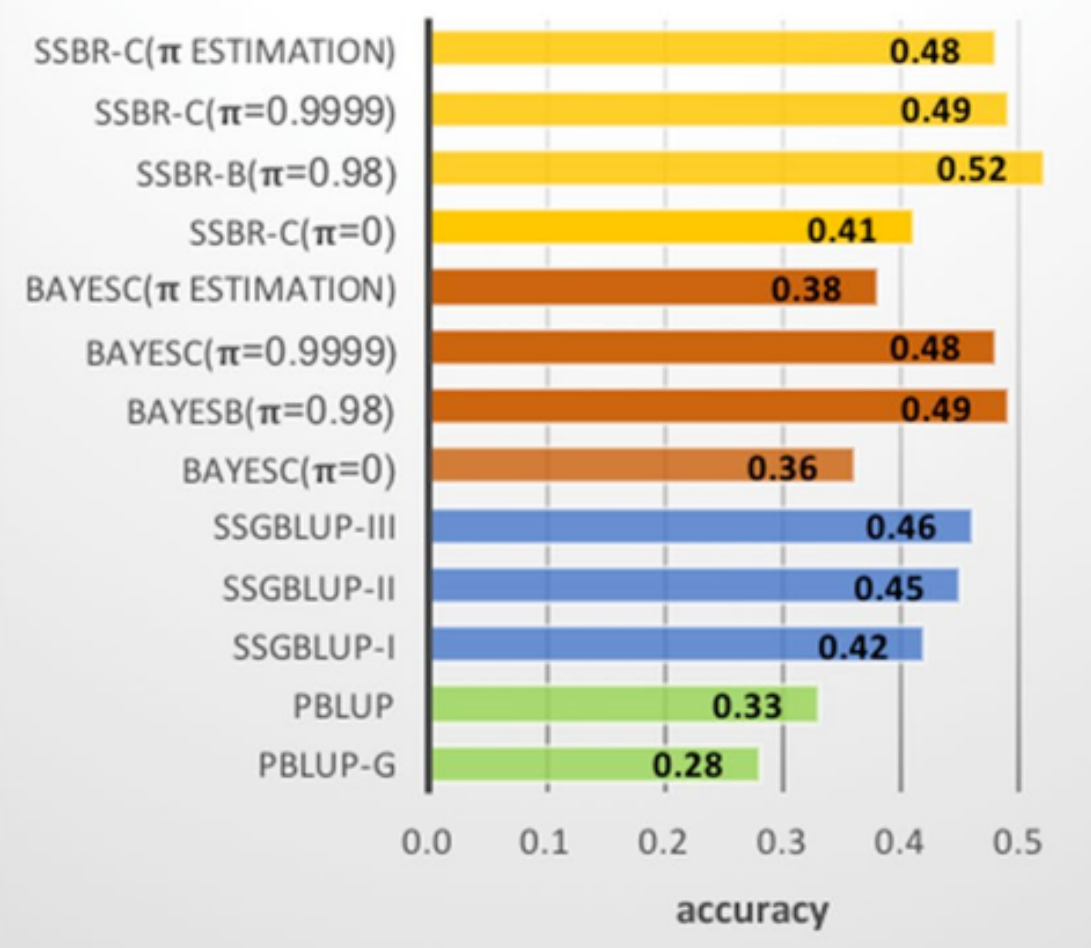
Hanwoo Eye Muscle Area



Non-genotyped data did not improve predictions

Lee *et al. Genet Sel Evol (2017) 49:2*

Hanwoo Carcass Weight



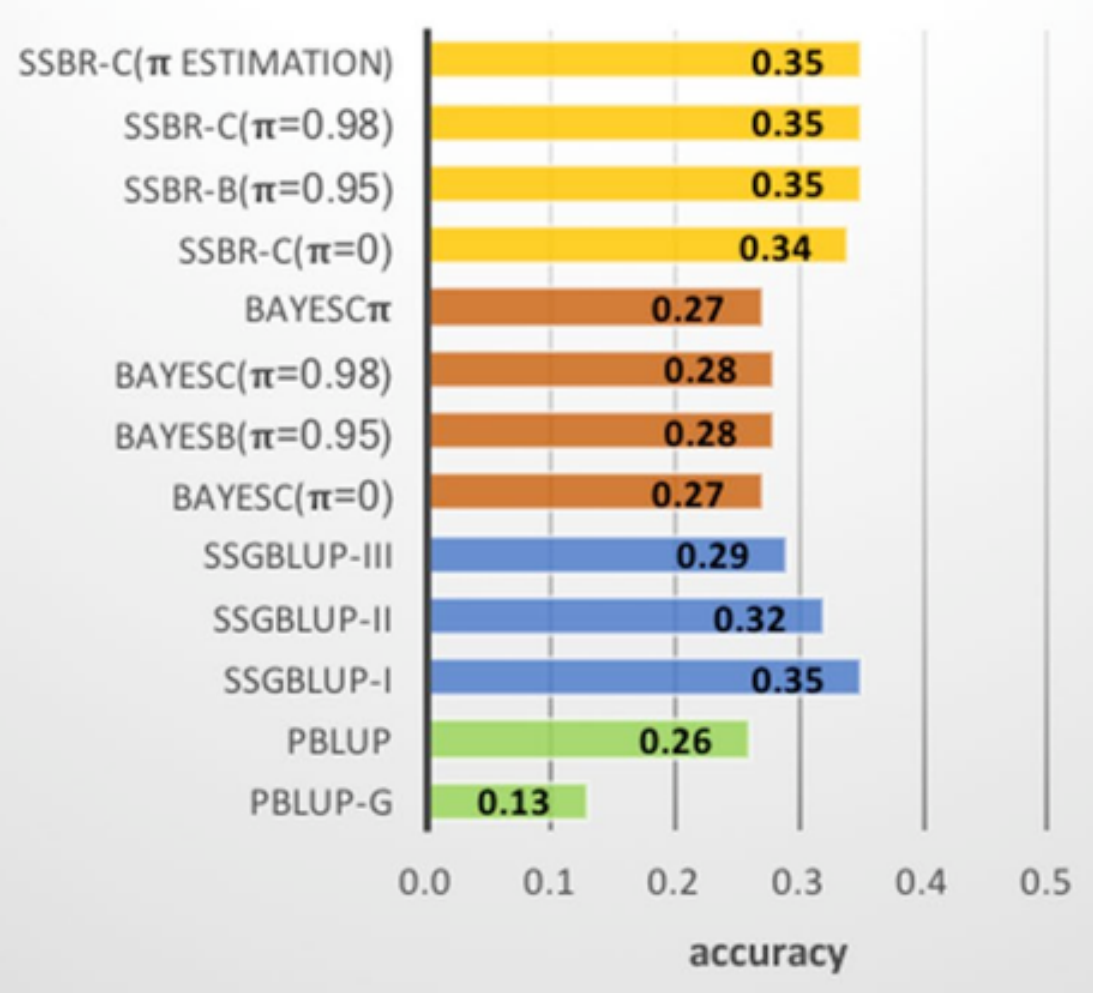
Major gene effects favoured variable selection models (ssBR)

Iterating ssGBLUP helped, but still not as good

Lee *et al. Genet Sel Evol (2017) 49:2*

ssGBLUP-II (and III) reweights marker effects from ssGBLUP-I)and II)

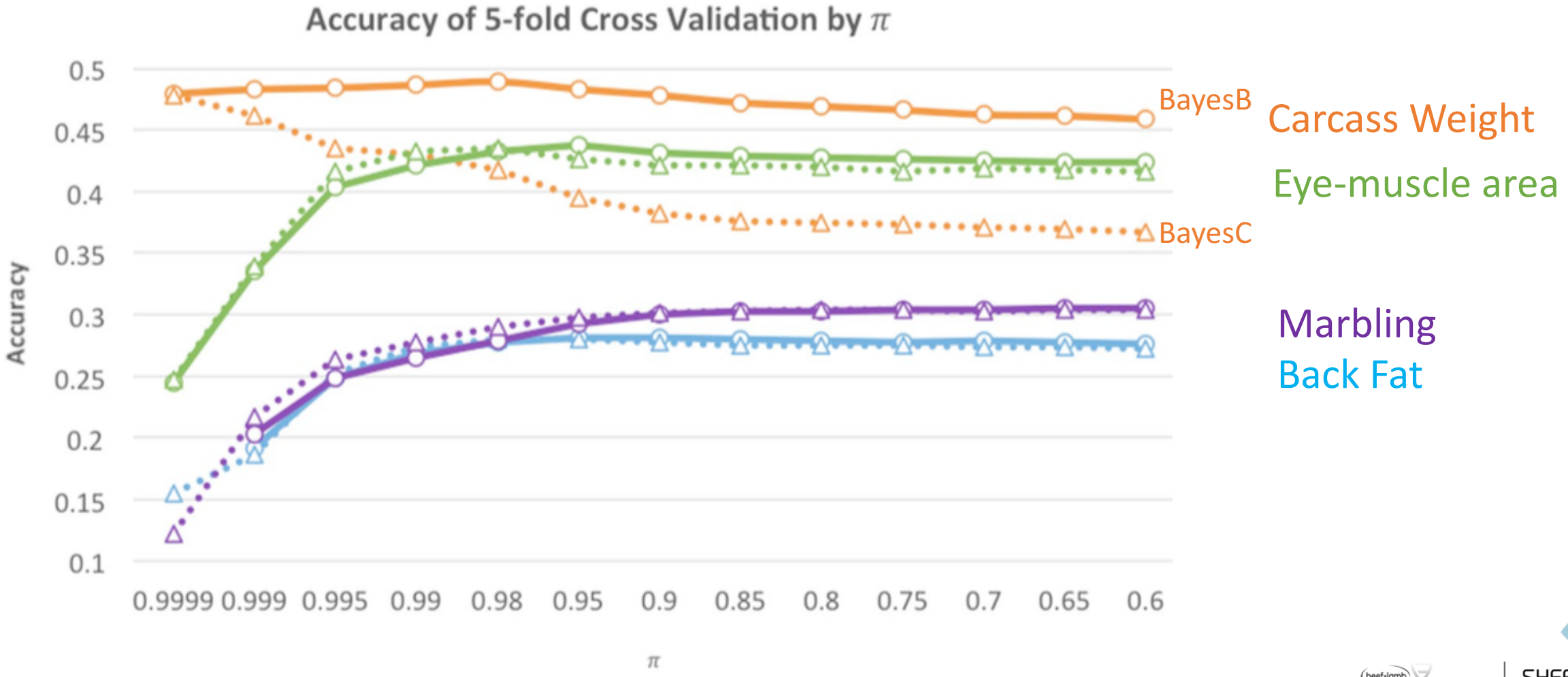
Hanwoo Back Fat



Iterating ssGBLUP made predictions worse

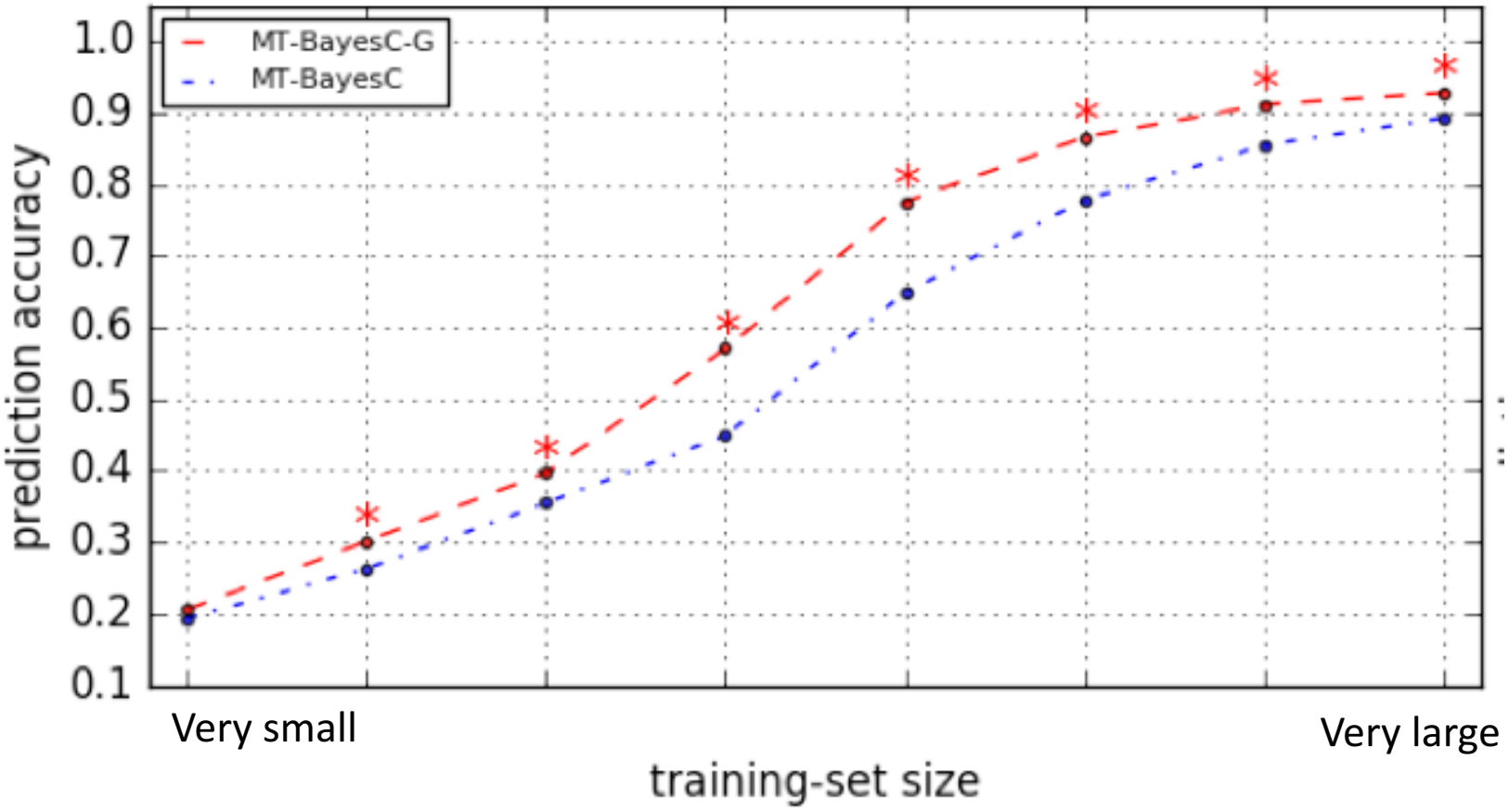
Lee *et al. Genet Sel Evol (2017) 49:2*

Genetic Architecture Differs Between Traits



beef+lamb GENETICS | SHEEP BREEDER FORUM

Multitrait Single-Step with different markers for each trait



Different markers for each trait

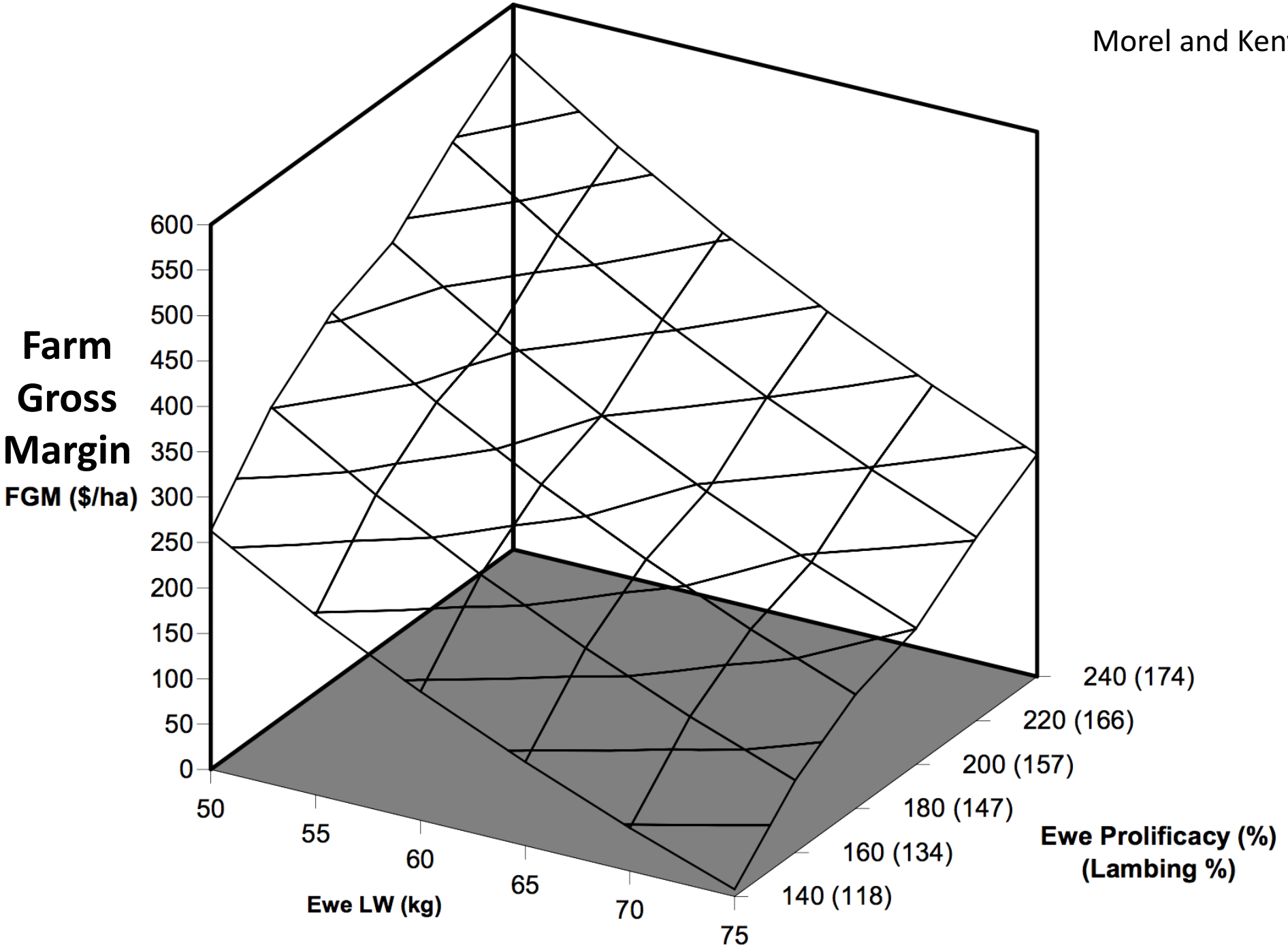
Same markers all traits

Simulated data

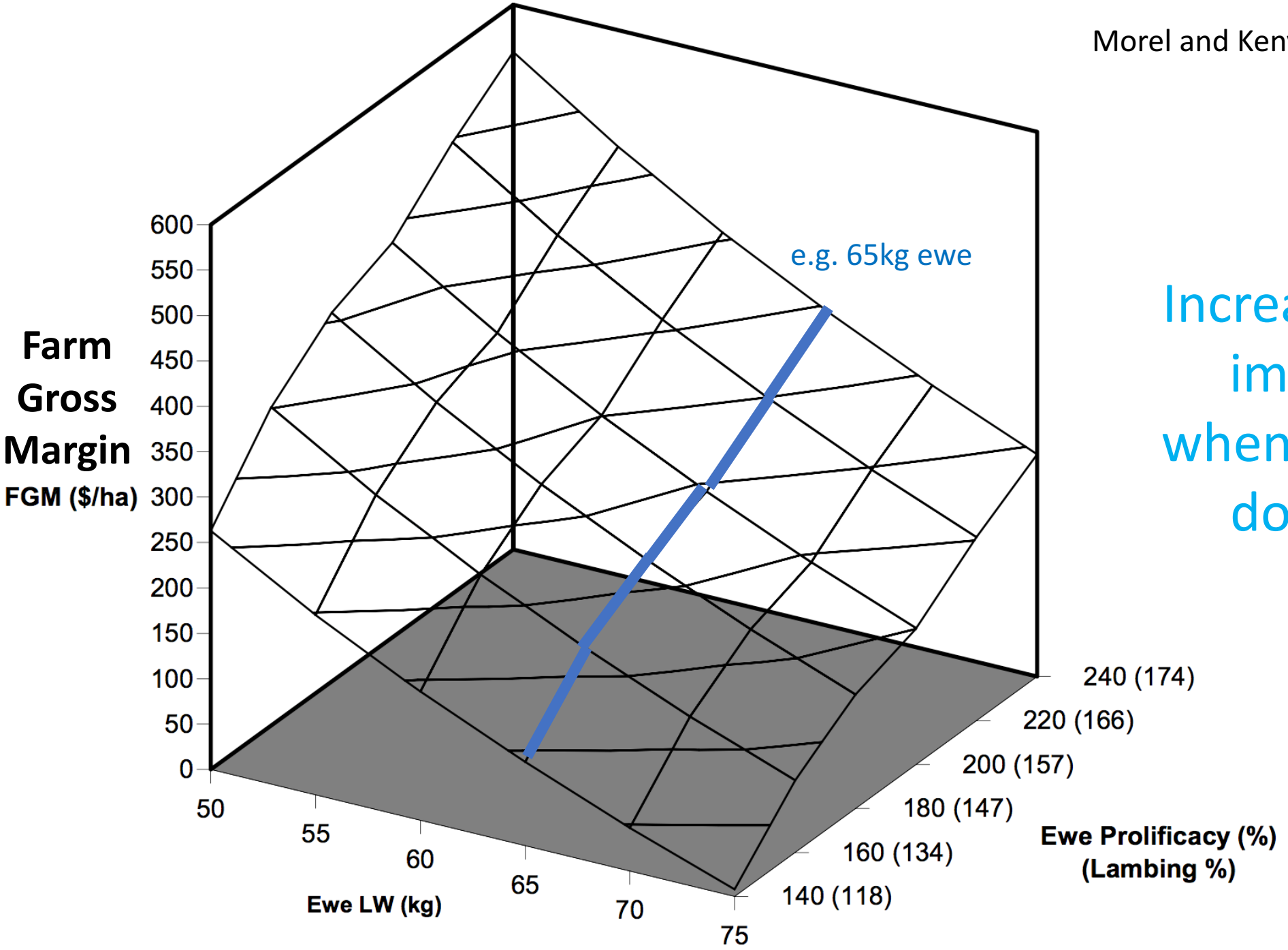
Cheng *et al* (2017) Under Review

Need to exploit extra accuracy with sensible selection

Morel and Kenyon (2006) NZSAP

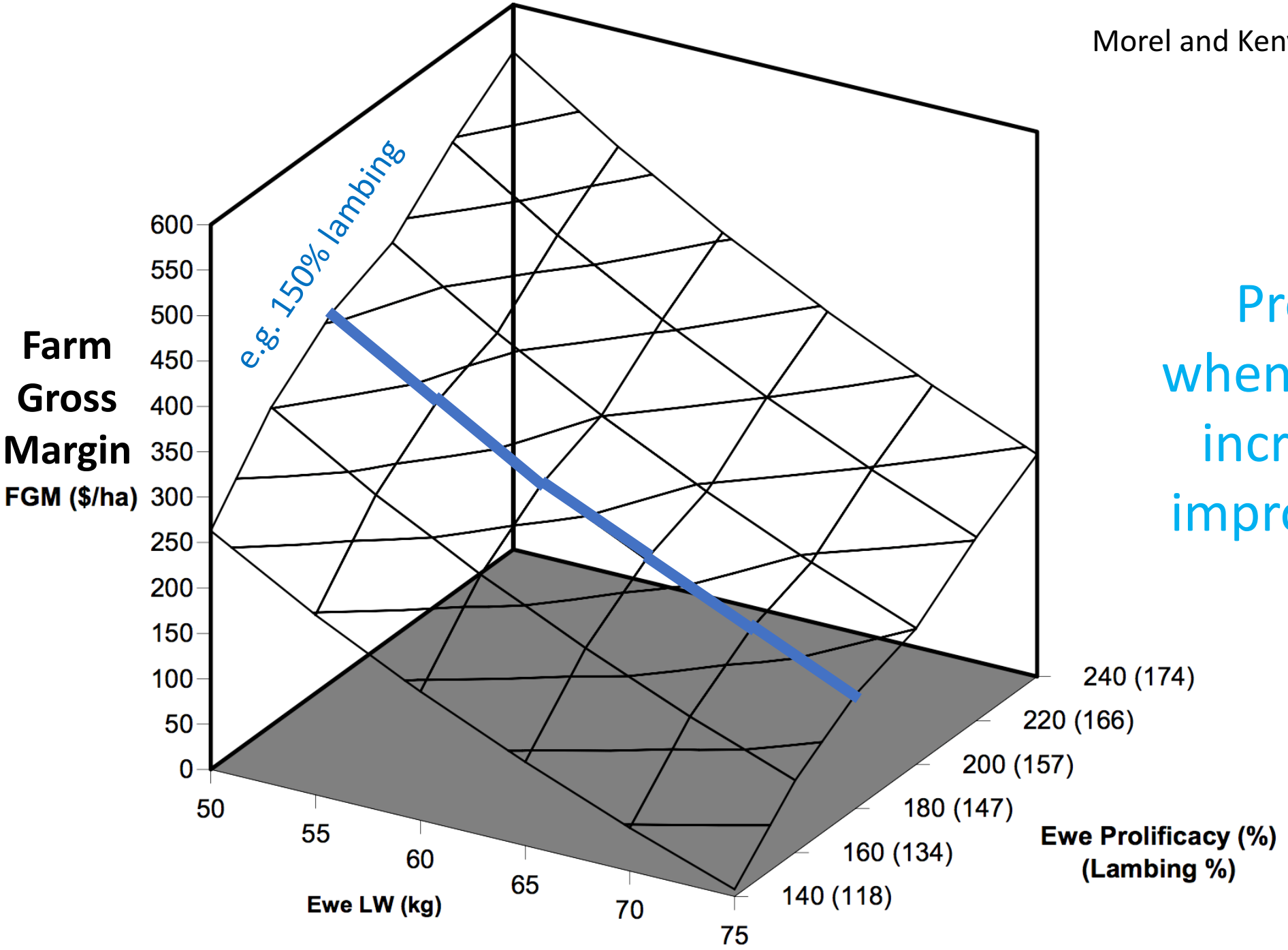


Morel and Kenyon (2006) NZSAP



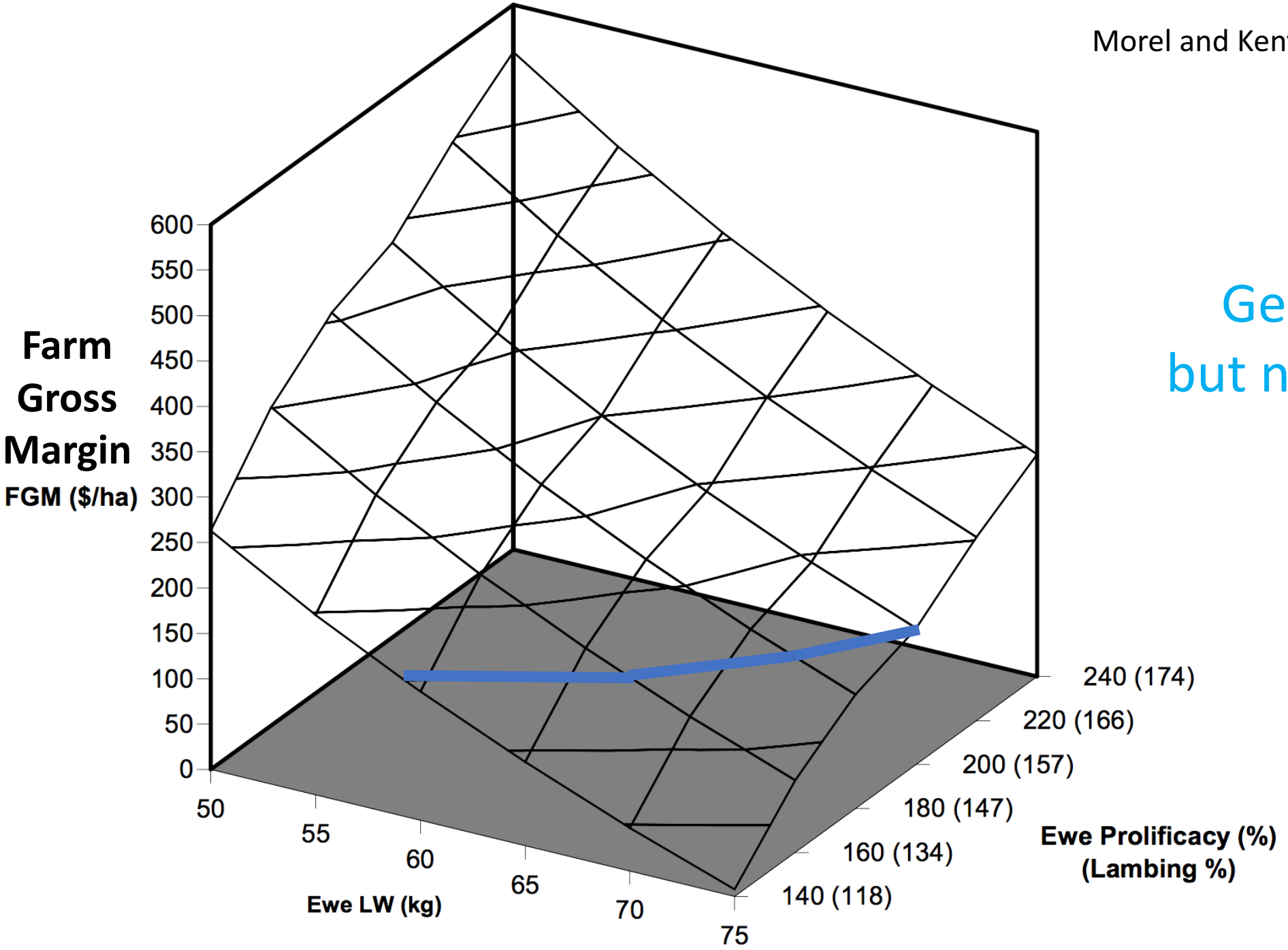
Increased prolificacy improves profit when ewe liveweight doesn't change

Morel and Kenyon (2006) NZSAP



Profit declines when ewe liveweight increases without improved prolificacy

Morel and Kenyon (2006) NZSAP

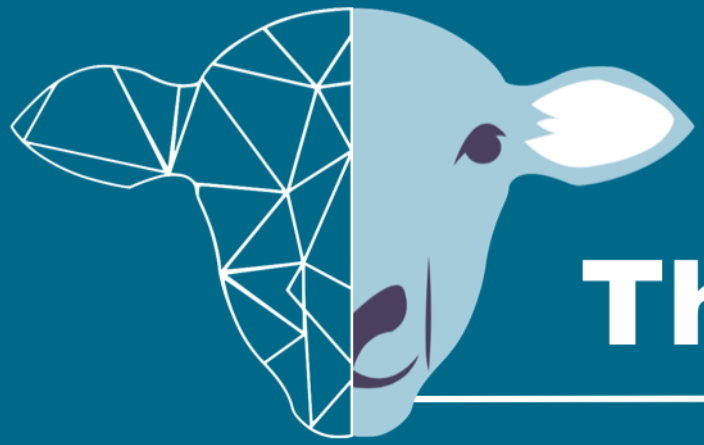


Genetic change but no improvement

60 kg ewe 118%
65 kg ewe 131%
70 kg ewe 144%
75 kg ewe 157%

Summary

- Genomic prediction does add costs as phenotypes are still required but it does add accuracy that can accelerate genetic gain
  - The benefits vary from trait to trait
    - Best for traits not measurable on the individual by selection age
  - The best value proposition is when it is used for nucleus ram selection rather than for evaluating sale rams
  - Greatest value from genomics may require different breeding schemes
- Single-step methods are preferred over two-step methods as they use all the data in the same analysis



Thank you.