Application of various models for the genomic evaluation of bovine tuberculosis in dairy cattle

R. Mrode, G. Banos, M. Winters and M. Coffey

Introduction

- Bovine tuberculosis (bTB) is a chronic bacterial disease of cattle caused *by Mycobacterium bovis*
- It presents a significant challenge to the UK cattle sector incurring annual costs of about £175 million
- Routine genetic evaluation for resistance to bTB has been implemented in the UK since January 2016.
- Trait was defined as positive skin test plus no positive skin test but having positive post–mortem examination results with infection rate of 8.29%
- bTB has a low heritability of about 0.09 and with bulls having an average reliability of 0.45

Objectives

- Study investigated whether inclusion of genotypic data might help increase accuracy.
- Some peculiar issues
  - Rate of infection is different for older vs younger bulls due to exposure time of their progeny to the disease.
  - With an all-or-none trait this can result in quite big shifts from one run to the next when progeny groups are still small.
  - Therefore validation candidates based on year of birth might not be optimum

- The study therefore looked at various models in addition to different methods of creating validation data sets
  - SNPBLUP, BayesCpi, Single-Step (SS)
  - Different levels of polygenic effects

Objectives

- Generally, polygenic effects are fitted to capture genetic variance not accounted by SNPs.
- Questions is; does including polygenic effects a uniform effect on SNPs of different allele frequencies?

- Therefore the impact of different levels of polygenic effects on SNP solutions for SNPs of different alleles is also examined.

Data for SNP-BLUP and BayesCpi

- Data consisted of 2232 Holstein-Friesian bulls
  - with deregressed proofs with at least 10 daughters and 40% reliability
  - Genotypes equivalent to the 50K chip were used - low density chips were all imputed to 50K chip and relevant SNPs extracted from HD chips
  - 43143 SNPs were analysed after edits
  - 1695 reference bulls were those born before 2007
  - 537 validation bulls born 2007 and onwards

Distribution of REF bulls and VAL by reliabilities

	REF	VAL
<=45	0	17
46-50	14	38
51-55	55	88
56-60	103	105
61-65	164	108
66-70	146	82
71-75	158	47
76-80	126	14
81-85	161	17
86-90	206	13
91-95	248	7
94-99	314	1

Different Validation sets

- Two additional different validation bulls were created
  - Random sample of first 30 bulls with reliability >=89 in the reference set plus all validation bulls with the same level of reliability (1888 bulls in REF & 344 in VAL)
  - Random sample of first 30 bulls with reliability >=93 in the reference set plus all validation bulls with the same level of reliability (2018 bulls in REF & 214 in VAL)

SNP-BLUP -Model and Analysis

SRUC

- Linear model consisting of
  - mean effect
  - random residual polygenic effect (0, 10, 20 and 30%)
  - random SNP effects

- Y- variable de-regressed sire proofs
- The number of daughters used as weights

- Accuracy were computed from correlations between DGVs and de-regressed proofs in validation set.

- BayesCpi - same model but with no polygenic effect

- Chain length was 80,000 with 24,000 regarded as burn-in period

Single step Analysis

- Analysis based on 607,929 cows with 934,987 records and a pedigree of 7,486034 animals
- Model of described in detail in Banos et al 2016 was fitted.
- Briefly an animal model
  - Fixed effects: mean, breakdown, year by month of breakdown, parity
  - Covariates: duration, age, %Holstein genes

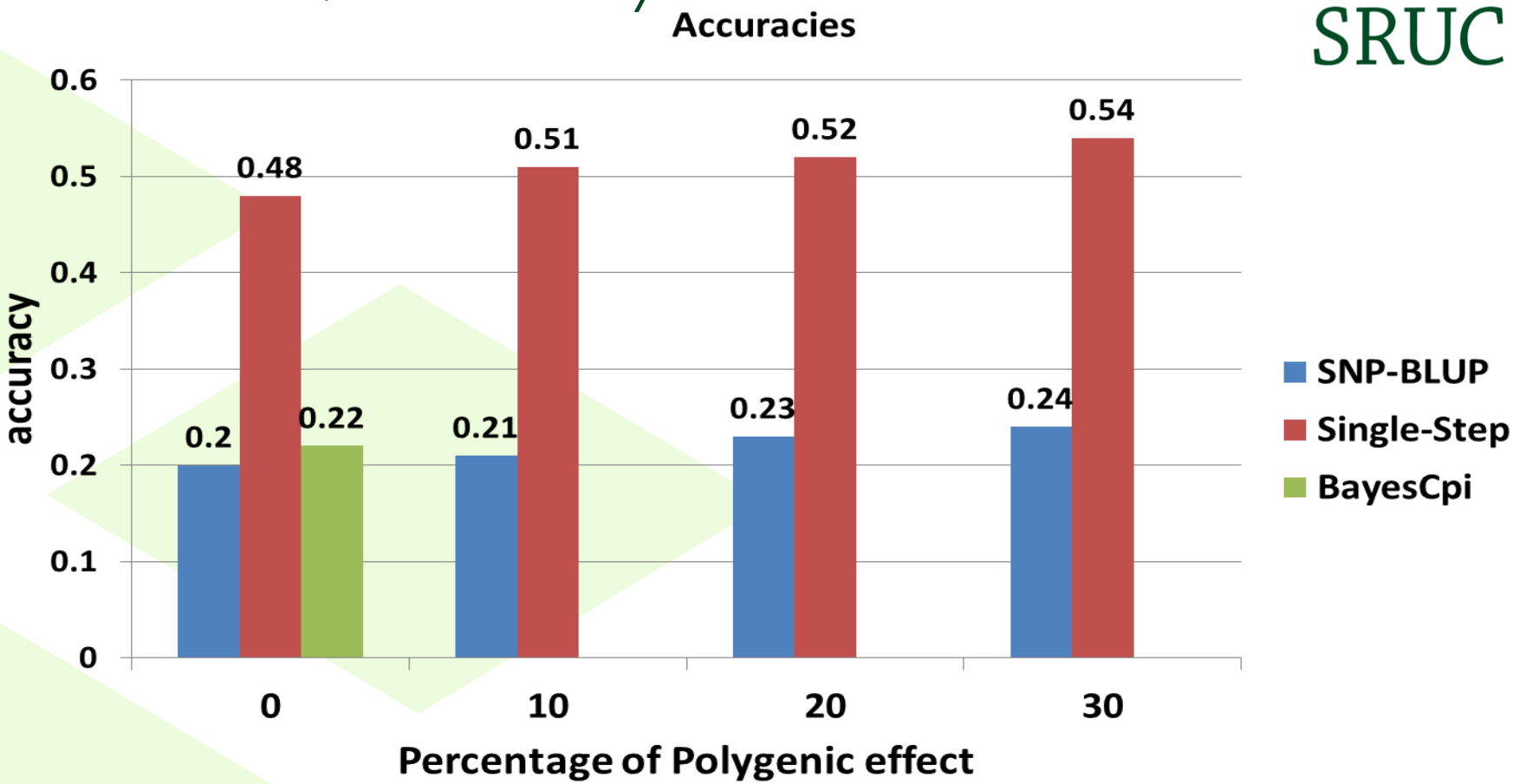
Single step Analysis

- 5435 sires of cows with records had genotypes
- **G** was computed for these sires
- The G_{22} matrix was then computed as
  $G_{22} = (1-w)G + wA_{22}$, with w set at 0, 10, 20 and 30%.

- The H^{-1} was then computed for all animals incorporating the G_{22} for the genotype animals.

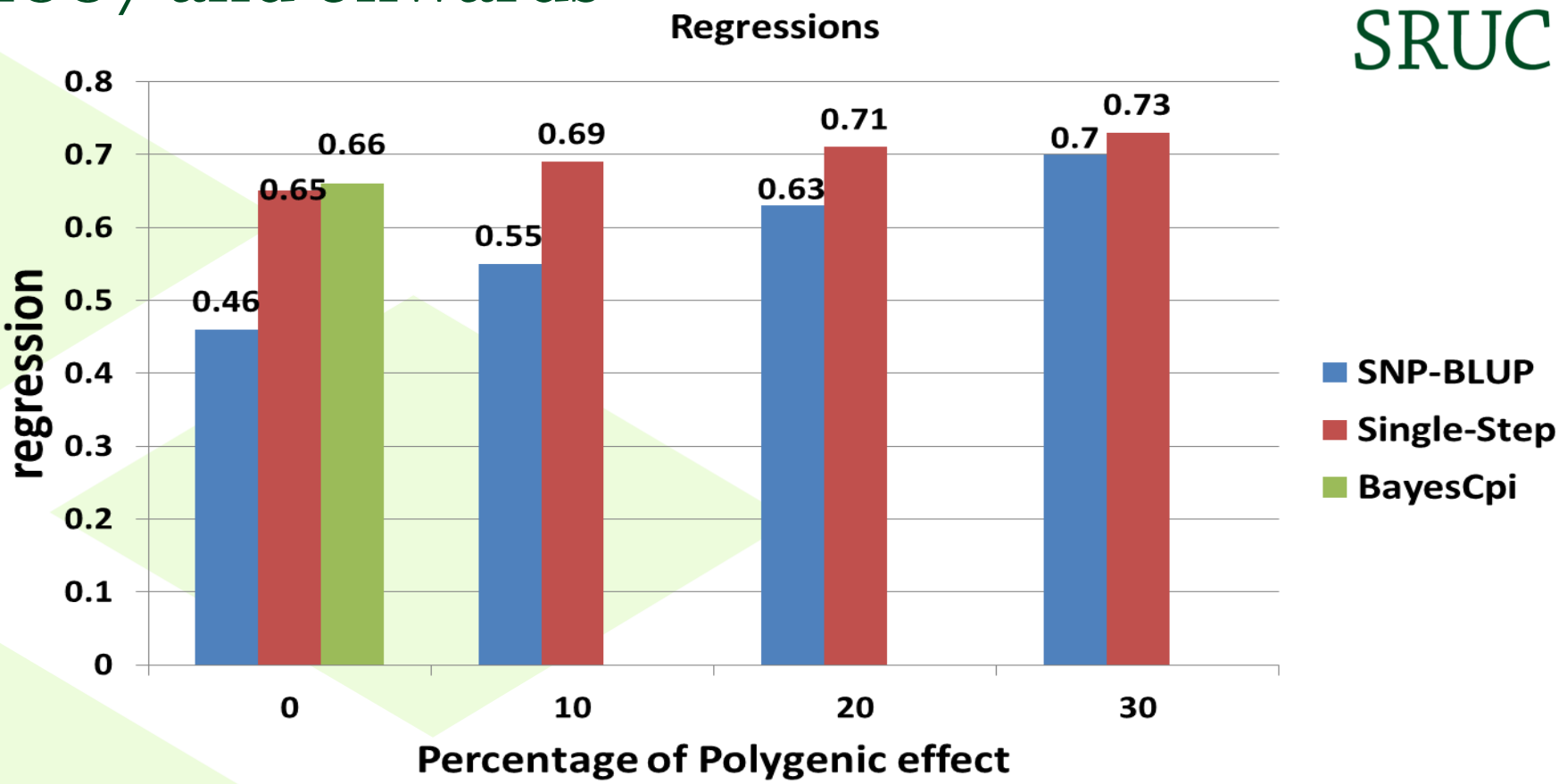
- The same set of validation bulls were also with records for their daughters set missing
- Accuracy were computed from correlations between GEBVs and
  - mean of the bull's individual daughter deviations
  - or de-regressed proofs in validation set

Accuracies of genomic prediction from validation bulls 2007 and onwards



- Corrections based on de-regressed proofs for Single-Step varied from 0.56 to 0.62

Regressions based validation bulls 2007 and onwards



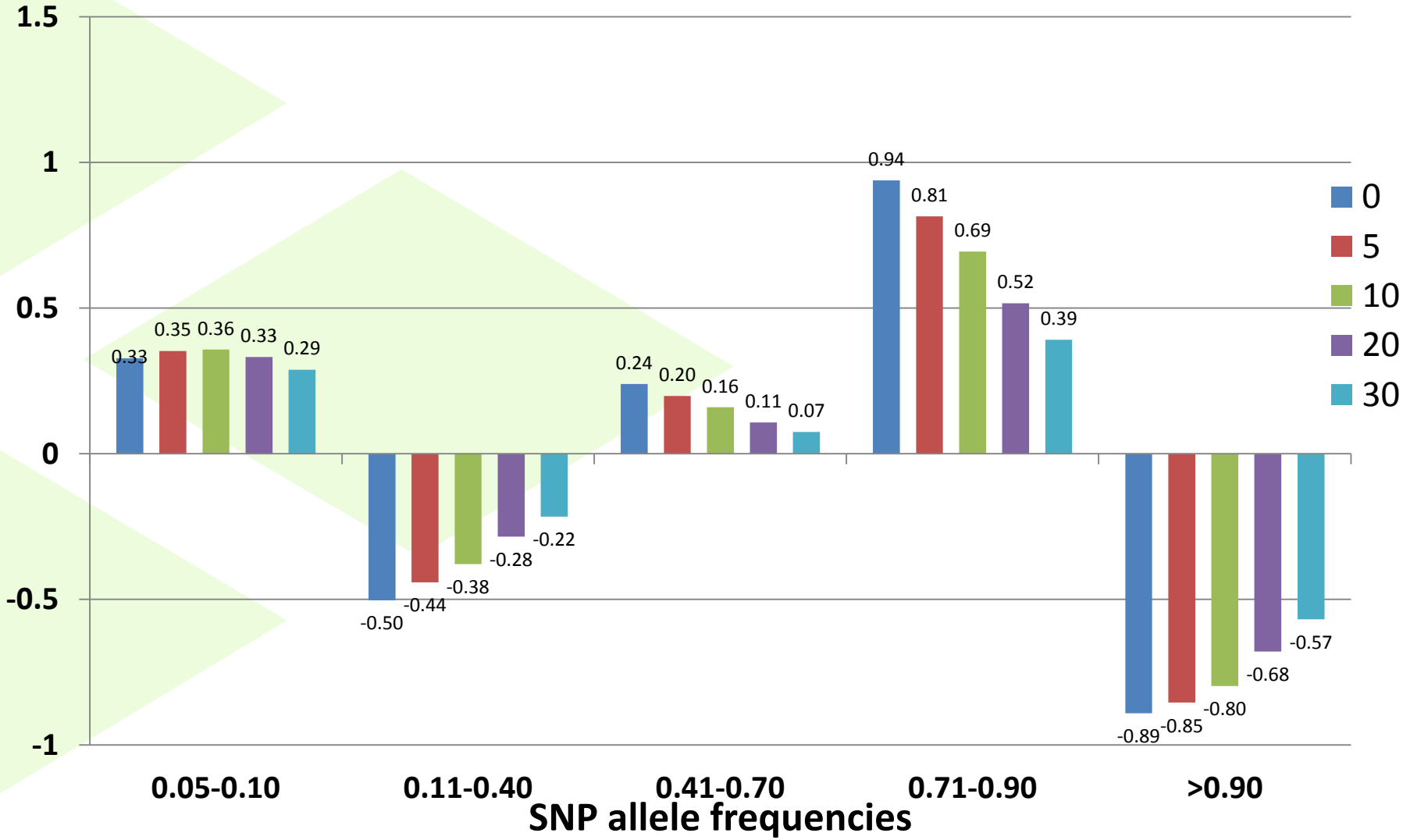
- Regression based on de-regressed proofs for Single-Step varied from 0.75 to 0.83

Results from alternative validation sets

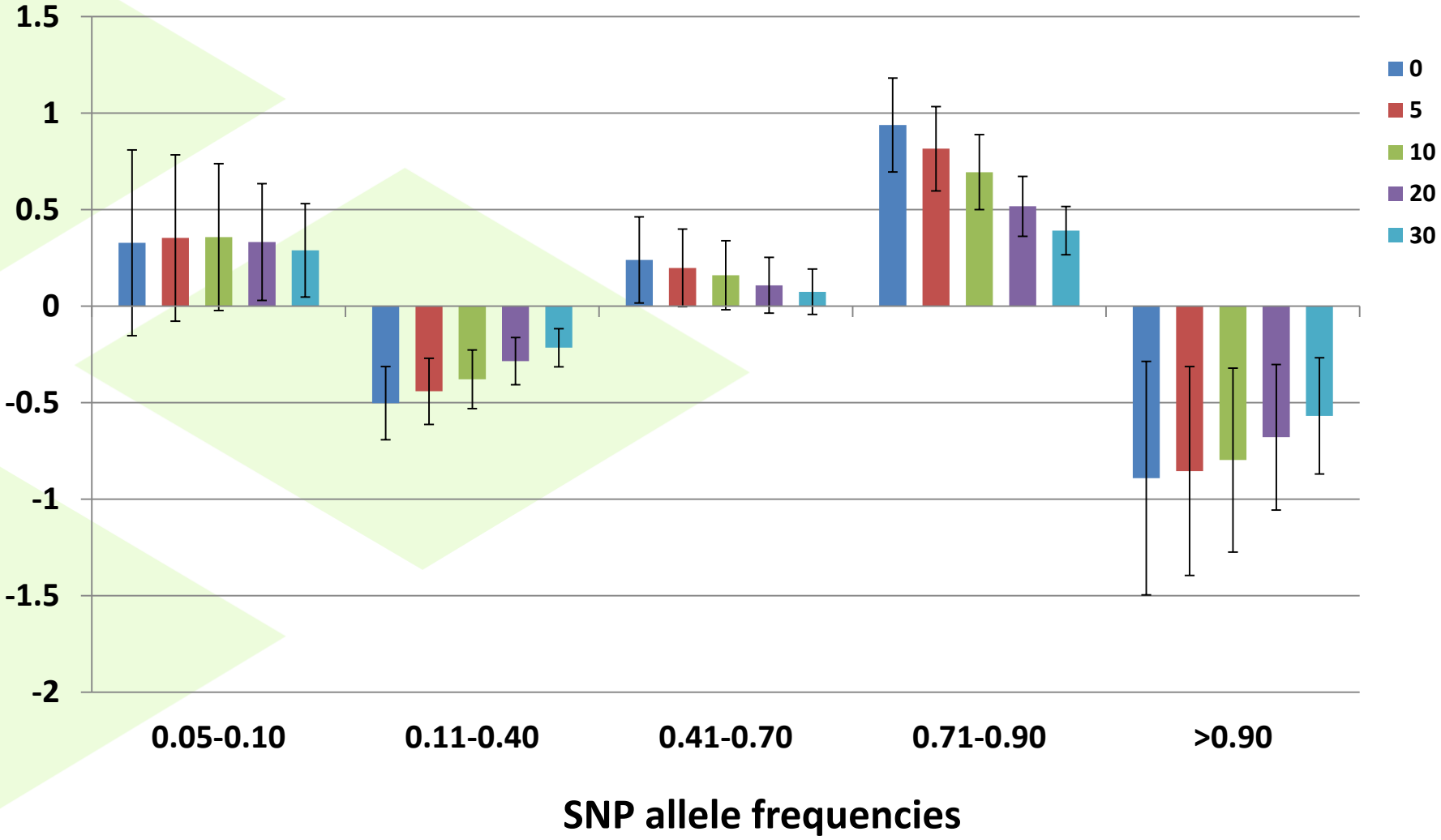
	SNP-BLUP (30% polygenic)		BayesCpi		Single-Step	
	Corr.	Reg.	Corr.	Reg.	Corr.	Reg.
>=89 Rel bulls	0.32	0.65	0.34	0.70	0.51	0.53
>=93 Rel Bulls	0.41	0.77	0.42	0.83	0.56	0.54

- Average accuracy for young animal from Animal model evaluations = 0.37

Mean SNPs effects at different levels of polygenic effects from SNP-BLUP



Mean SNPs effects with SEs at different levels of polygenic effects from SNP-BLUP



Conclusions

- Given the data structure and size
  - Single-Step evaluations seems the most appropriate to apply in this study
- Definition of validation data sets to capture similar rate of infection as in the reference sets seems crucial for SNP-BLUP & BayesCpi
- Incorporating genotypes information resulted in increased accuracies
- Fitting a polygenic effect does not have a uniform impact on the estimates of SNP effects
  - Its influence is dependent on the allele frequency of the SNP

Acknowlegements

- Funding by AHDB Dairy gratefully acknowledged

ICAR 2017 UK TECHNICAL WORKSHOP

ICAR THE GLOBAL STANDARD FOR LIVESTOCK DATA

ICAR CONFERENCE 2017

14–16 June

Edinburgh, Scotland

To register your place at ICAR 2017 visit www.icar2017.co.uk.

Early bird prices available until March 2017.

SRUC

AHDB DAIRY