Genomic evaluations of sheep in New Zealand

MA Lee[a,b*], S-AN Newman[c], KG Dodds[c], MA Nilforooshan[b,e], B Auvray[d], S McIntyre[a] and D Campbell[a]

[a]*Beef+Lamb NZ Genetics, PO Box 5501, Dunedin 9054, New Zealand;* [b]*Department of Mathematics and Statistics, University of Otago, PO Box 56, Dunedin, New Zealand;* [c]*AgResearch, Invermay Agricultural Centre, Private Bag 50034, Mosgiel, New Zealand;* [d]*Iris Data Science, Level 2, 7 Bond St, Dunedin, New Zealand; Livestock Improvement Limited, 605 Ruakura Road, Newstead 3286, New Zealand*

[*]*Corresponding author:* Email: michael.lee@otago.ac.nz

Abstract

Genetic improvement of farmed livestock in New Zealand is achieved through selective breeding using predictions of an animal's genetic merit called 'breeding values'. Inclusion of genomic data in genetic evaluations (known as 'genomic selection') has increased the accuracy of breeding values compared to best linear unbiased prediction (BLUP) methods based on relationships from known pedigree; particularly for young animals with limited recording. The additional accuracy gained from genotype information is most valuable for traits that are: hard to measure; sex limited; measured late in life; and/or expensive to measure. The 'New Zealand Genetic Evaluation' (NZGE) for sheep has used single-step genomic BLUP (ssGBLUP) since late 2018, which provides the basis for genomic selection in the NZ sheep industry. This method allowed pedigree, performance records and genomic data to be included in the same analysis to predict breeding values in contrast to performing a separate pedigree and genomic analysis and then combining these results (multi-step genomic BLUP; msGBLUP). ssGBLUP is a better alternative for genomic evaluations compared to msGBLUP from a perspective of bias, best use of information and computational-ease. This paper describes the NZGE genomic analysis and details some of the current challenges relating to optimising accuracy and minimising bias in the evaluation. Genomic selection is a useful additional tool for breeders. However, it is important to be mindful that data recording is still the foundation for any genetic improvement program, as is allowing breeders to best leverage genomic selection.

Keywords: sheep; genetics; genomics; breeding

Introduction

The best linear unbiased prediction (BLUP) statistical method (Henderson et al. 1959; Henderson 1975) has been used to predict breeding values in sheep in New Zealand for decades (Newman et al. 2000). In BLUP (equation 1), the breeding values (**u**) are predicted from the information from phenotypes (**y**) and fixed effects (**b**), where **X** and **Z** are design matrices and **e** is a vector of residual errors.

Equation 1

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$$

In order to predict breeding values or random effects ($\hat{u}$) from equation 1, the mixed-model equations (equation 2; MME in matrix notation) described by Henderson 1975, where λ is the ratio of the variances of the residual additive effects over the variances of the additive genetic effects assuming an animal model are solved. The model assumes homogeneous variances for the residual errors (e) that are independent of the random effects ($\hat{u}$). This also implies that population genetic parameters are known (e.g., genetic and residual additive variances), requiring pre-estimation of parameters for the population to be evaluated.

BLUP assumes an infinitesimal or polygenic model (Falconer 1996) – in that a given quantitative trait is explained by an infinite number of genes, each having a small effect.

The relationship among the breeding values is given by the covariance matrix (**G**) that is estimated from pedigree for BLUP, so that the numerator relationship matrix (A) is used for G in the MME. Increasingly, genotype data are used to estimate the relationship matrix. For example, for genomic BLUP (GBLUP) the genomic relationship (**G**), calculated from genotype data, is used in the MME.

Equation 2

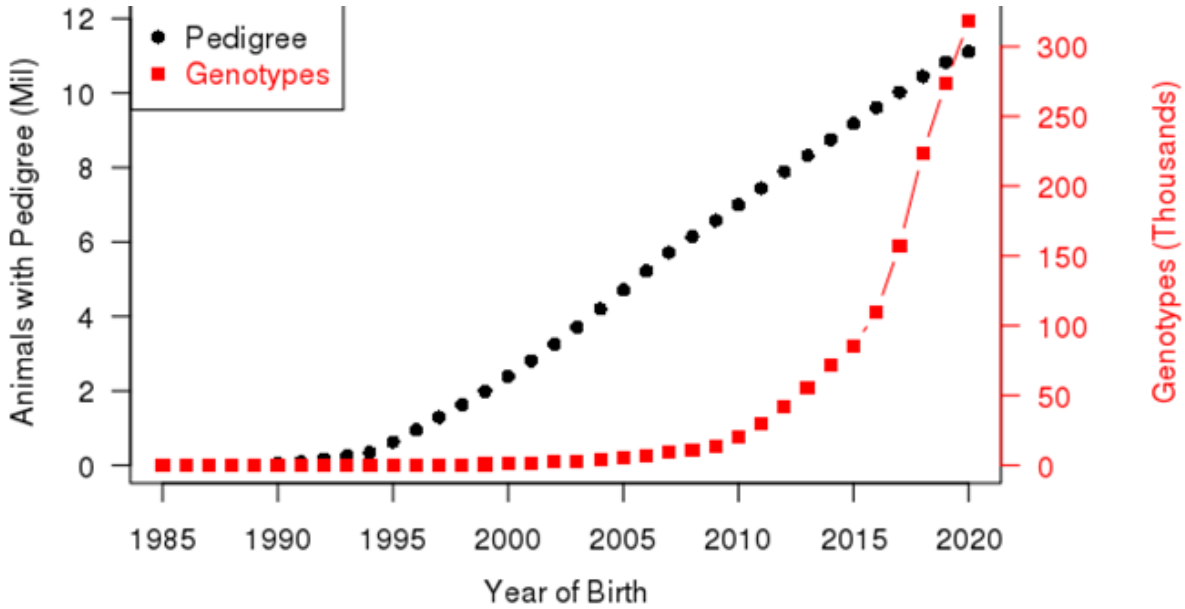
$$\begin{pmatrix} X'X & X'Z \\ Z'X & Z'Z + G^{-1}\lambda \end{pmatrix} \begin{pmatrix} \hat{b} \\ \hat{u} \end{pmatrix} = \begin{pmatrix} X'y \\ Z'y \end{pmatrix}$$

A national across-flock and -breed genetic evaluation has been available to the NZ sheep industry through Sheep Improvement Limited (SIL) called the New Zealand Genetic Evaluation (NZGE). Genomic data have been incorporated into the NZGE, but the models and genetic parameters that have been determined over the past decades and used in SIL BLUP analyses still form the basis for NZGE, irrespective of the integration of genomics. Most the traits are analysed in multi-trait models that use genetically correlated traits to improve accuracy and account for selection bias. NZGE is computed using cloud-based high performance computing infrastructure and reported weekly to suit the structure of sheep seed-stock breeders, which comprises many different breeders needing evaluation results at various times.

The number of pedigreed animals available for evaluations has increased from thousands of animals in the 1980s to a total of about 10.75 million in 2020; whilst the number of animals with genotypes (over 5,000 markers) has increased to over 300,000 (Figure 1).

Figure 1 Cumulative number of sheep evaluated in NZGE with pedigree (in millions (Mil)) and genotypes

Figure 1 Cumulative number of sheep evaluated in NZGE with pedigree (in millions (Mil)) and genotypes from 1985-2020 calculated by the year of birth of animals.



from 1985-2020 calculated by the year of birth of animals.

This paper describes the NZGE with a focus on the implementation of the genomic component and highlights some of the challenges.

Earlier genomic predictions via GBLUP

The basis of genomic prediction for sheep was made possible after the development of the Illumina Ovine 50K single nucleotide polymorphism (SNP) chip as part of the International Sheep Genomics Consortium in 2009, comprising AgResearch, Baylor College of Medicine (Houston, TX), the Commonwealth Scientific and Industrial Research Organisation (Australia), and the United States Department of Agriculture. Since then, a number of SNP chips have been developed and used. The density ranged from an early chip of about 5K to a high density 600K chip. Significant numbers of animals were only genotyped for genomic prediction once cost-effective, lower density, SNP chips (~5K) were available.

The first genomic predictions in 2013 were based on Genomic BLUP (GBLUP) (Meuwissen 2009), where the genotypes were used to create a genomic relationship matrix (**G**) that is used in the MME instead of the pedigree relationship matrix (**A**) (equation 2). The process was described by Auvray et al. 2014 and predictions were restricted to the maternal breeds Romney, Coopworth, Perendale and composites of these breeds.

A problem with GBLUP is that breeding values are predicted only for genotyped animals and most of the animals to be evaluated from a NZGE perspective have only pedigree data (see Figure 1). Consequently, the strategy for genomic prediction was to prioritise sires for genotyping and to optimally use all information by calculating a pseudo-phenotype (e.g. deregressed breeding value with parent average removed) for these animals. These animals formed the reference population which were used to calibrate the genomic predictions. However, as these animals also had a breeding value from BLUP (estimated breeding value; EBV) there was a need to combine or blend the predictions from GBLUP (molecular breeding value; MBV) with EBV (equation 3). The blending function was based on a selection index approach (VanRaden 2008; Harris & Johnson 2010), where the blended genomic breeding value ($\hat{g}$) is calculated from three different evaluations namely, the GBLUP and two pedigree BLUP evaluations, consisting of a "national" evaluation and a subset evaluation of only genotyped animals. The reliabilities from these three evaluations were, respectively R_G, R_N and R_A and the breeding values to be blended, respectively, value $\hat{a}_G$, $\hat{a}_N$ and $\hat{a}_A$. Importantly, the $\hat{a}_G$ and $\hat{a}_N$ are not assumed to be independent and their assumed covariance is calculated from reliabilities of the breeding values by equation 4.

Equation 3 Blending of EBV and GBV

$$\hat{g} = \frac{(1-R_S)(1-R_G)\hat{u_N} - 1 - R_N 1 - R_G \hat{u_S} + (1-R_N)(1-R_S)\hat{u_G}}{1 - 2R_S - R_N R_G + R_S R_G + R_N R_S}$$

Equation 4 Covariance of the national and molecular breeding values

$$cov(\hat{u_G}, \hat{u_N}) = R_S + \frac{(R_G - R_S)(R_N - R_S)}{1 - R_S}$$

The number of genotyped animals was modest enough to estimate the accuracies of the breeding values from the solutions of the left hand side of the MMEs. This allowed accuracies to be calculated for each animal to be evaluated and used for blending in equation 3 as a reliability. Similarly, when the software ASREML (Gilmore et al. 2009) was used for BLUP the accuracies could also be calculated. The goal of the blending procedure was to allow, as much

as possible, breeders to have one breeding value that was comparable across groups of animals, such as flocks.

However, this multistep evaluation was problematic as recalibrations (i.e. including new information into the components needed to run GBLUP) were time-consuming and the use of all available information was not optimal, and with each calibration there was significant re-ranking amongst animals.

The New Zealand Genetic Evaluation

Single-step GBLUP

NZGE with ssGBLUP (Aguilar et al. 2011) was implemented in 2018 to allow for a single national evaluation that included genomic information and in which most animals still have only pedigree without any genotype data. This replaced the pedigree BLUP NZGE that was the main national evaluation for sheep in NZ at that time.

ssGBLUP was attractive as the models (mostly multi-trait), genetic parameters and much of the infrastructure, developed over a number of decades, to adjust and quality control the data could be re-used and the transition would allow for minimal changes (re-ranking), particularly for those animals without genomic information.

In principle, ssGBLUP is simple in that the relationship matrix used in the MME is replaced by a matrix constructed from pedigree and genotypes. The pedigree matrix (**A**) can be partitioned as in equation 5, where the partition $\mathbf{A_{22}}$ corresponds to animals with genotype and $\mathbf{A_{11}}$, $\mathbf{A_{12}}$ **and** $\mathbf{A_{21}}$ are those without genotypes. The combined pedigree and genomic relationship matrix (**H**) used in ssGBLUP is given in equation 6. Importantly, **H** gives the covariances of all animals to be evaluated using both genotype and pedigree information.

Equation 5

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

Equation 6

$$H = \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} + A_{12}A_{22}^{-1}GA_{22}^{-1}A_{21} & A_{12}A_{22}^{-1}G \\ GA_{22}^{-1}A_{21} & G \end{pmatrix}$$

In practise, the $\mathbf{H^{-1}}$ is calculated directly as in equation 7 and used in the MME ($\mathbf{G^{-1}}$ in equation 2).

Equation 7 Calculation of the H inverse.

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

G generally needs to be regularised (Mäntysaari et al. 2017) before solving and in NZGE it is weighted with $\mathbf{A_{22}}$ (**A** in equation 8) by a value of α to give G_w. However, the value of α is also optimised within trait by validation. An interpretation of α is the proportion of additive genetic variance that is not explained by **G** and in NZGE this can vary from 0.05-1.

Equation 8 Weighting of the **G** matrix with **A** ($\mathbf{A_{22}}$).

$$G_w = (1 - \alpha)G + \alpha A$$

As an example, a validation for fleece-weight at 12-months of age (FW12), within a multivariate model that included weigh and lamb fleece weight, using ~10 million pedigreed animals and 50,630 genotyped animals each with 41,958 SNPs to make **G**. The genotypes were deduced from a 600K, 50K or lower than 50K chip. The lower density genotypes were imputed to 41,958 SNPs using FImpute2 (Sargolzaei et al. 2014), where only genotypes that imputed with an accuracy >97% concordance were used. The accuracy of imputation was determined by validation within the chip used.

In order to estimate an optimal α, 810 progeny-tested sires (born from 2013-17) were used as validation animals. The mean number of progeny with phenotype for the sires was 38.6 (range 21-69). A reduced dataset was constructed by removing all the phenotypes from the sires and all their descendants. Deregressed breeding values with parent average removed (dBV) (Garrick et al. 2009) that were calculated from a BLUP using the full dataset with 2,016,945 phenotypes were used to compare against predicted breeding values with the reduced dataset to assess the prediction accuracy of ssGBLUP and pedigree BLUP. The results, analysed within and across breed groups, as correlation and the slope from regressing dBV on predicted breeding values are given in Table 1. The results from a Romney flock that was sizeable and well-recorded with a significant amount of genotyping are also given in Table 1, where 67 of the 810 sires used in validation were from this flock.

In general, ssGBLUP correlations were higher than for BLUP correlation across all α-values and groups. The slopes from regression also suggested ssGBLUP was also less biased than BLUP. The correlation was 0.26 for BLUP and at best 0.44 for ssGBLUP which was also less biased. In NZGE an α-value of 0.25 is used. The aim is to use a value that optimises accuracy and minimises bias across all breeds. A similar procedure is used to estimate α across all other NZGE traits.

Imputation

Most of the genotypes submitted for genomic predictions are from lower density chips than 50K, for economic reasons. These genotypes are imputed to a density of about 42K and used for ssGBLUP. Imputation refers to the inference of unobserved genotypes (e.g., inferring the genotypes that are on the 42K panel of genotypes that are not on the lower density 15K panel). We recommend that breeders routinely genotype the sires that they use for mating with a higher-density SNP chip (e.g., 600K) to update imputation reference panels.

Imputation inaccuracy reduced prediction accuracy and could result in bias. To demonstrate this, we used a prediction dataset of 1,301 Romney animals that were genotyped with an Ovine 15K SNP chip and 15,885

Table 1 Correlations (r) and slopes from a validation for ssGBLUP with FW12. The slopes were the dBVs regressed onto the predictions for the animals used in validation.

	ssGBLUP alpha value used									
	1 (BLUP)		0.75		0.5		0.25		0.05	
Group	r	slope	r	slope	r	slope	r	slope	r	slope
All	0.49	0.88	0.53	0.88	0.54	0.85	0.54	0.8	0.5	0.7
Romney	0.43	0.85	0.46	0.89	0.47	0.84	0.46	0.76	0.41	0.62
Coopworth	0.31	0.6	0.36	0.67	0.4	0.7	0.43	0.7	0.41	0.63
Perendale	0.38	0.74	0.42	0.77	0.44	0.76	0.44	0.74	0.43	0.68
Composites	0.54	0.59	0.57	0.76	0.58	0.88	0.57	0.86	0.52	0.42
Romnney flock	0.26	0.59	0.38	0.76	0.42	0.76	0.44	0.7	0.42	0.59

suitable training animals genotyped with a 50K chip. The total number of pedigreed animals in the evaluation was 9,513,189. Different imputation reference panels, generated by random sampling from available reference genotypes, were used to impute the genotypes of the prediction dataset and the accuracy of imputation, for each animal for a given imputation reference panel, was estimated by validation.

An example for the trait number of lambs born is given in Figure 2, where the prediction dataset had an imputation accuracy of 89% (panel A; GBV89) and 70% (panel B; GBV70). The breeding values of animals imputed with an accuracy of 89% were indistinguishable from those that were imputed to a high average accuracy (96%) with a correlation of >0.99. Whereas those imputed to an average accuracy of 70% were less accurate, with a correlation of 0.9 compared to those imputed at 96%. They were also biased (in this case upwards), such that there was a shift in the mean for these animals. From the perspective of ram breeders, this bias might result in sub-optimal predictions across groups (e.g., flocks). Poor convergence in solving the ssGBLUP equations was another result of increased numbers of poorly imputed genotypes in the evaluations. We observed similar results for other traits (results not shown).

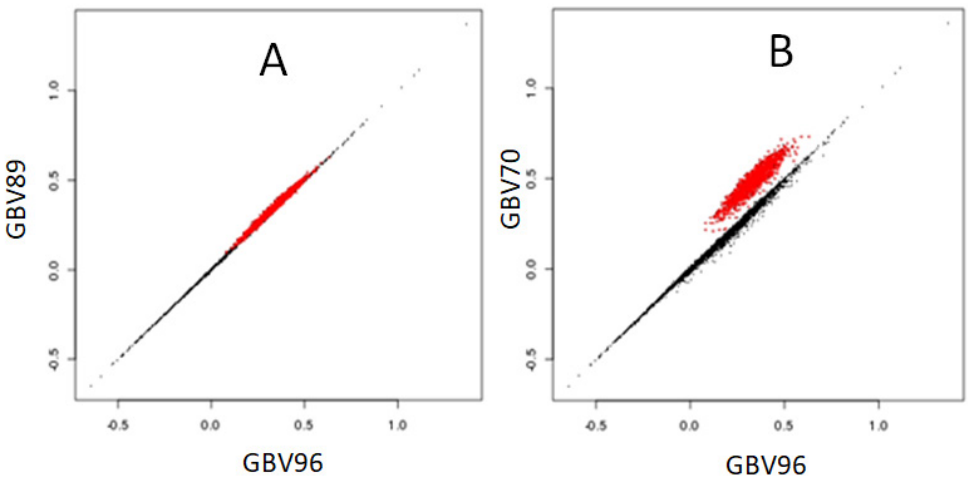
In practise, with NZGE the imputation accuracy of each animal is estimated by validation and only animals that reach a threshold of imputation accuracy (90%) are included. This is also a useful quality control check to ensure that only genotypes from animals that are likely to obtain useful predictions are included (i.e. correct breed and relatives of previously genotyped and phenotyped individuals). This is because the reference panel comprises animals of correct breed that are generally sires from flocks using genomic selection. Therefore, animals that are of the wrong breed type and unrelated will generally be poorly imputed.

Traits with low levels of recording

There are some traits in NZGE that have low levels of recording - often traits for which phenotypes are difficult and/or expensive to measure and where breeders are generally unwilling to record. As such, they are also the traits for which genomic predictions are most useful (e.g., faecal egg count (parasite resistance) and green-house gas emission (Pickering et al. 2015)), where the ratio of animals with phenotype to those with genotype is low.

As the phenotyping to genotyping ratio decreased for a trait, the convergence for ssGBLUP takes more iterations to

Figure 2 Graph of breeding values for the trait number of lambs born from a ssGBLUP with prediction dataset genotypes imputed to an average accuracy of 89% (GBV89) and 70% (GBV70) given in red compared to when imputed to an accuracy of 96% (GBV96) . Sheep that were not imputed (training animals) are given in black.



a point where there is a failure to converge. Consequently, in NZGE for a given trait, genotypes from animals that meet all the following criteria: are not sires, have no phenotype; were culled or deceased and with a recorded year of birth greater than two years in the current evaluation year are not included. These animals do not offer any useful information and are not of interest to breeders, but they remain in the evaluation as a pedigreed animal.

A longer-term solution to cater for breeding schemes and scenarios for groups that undertake limited or no phenotyping or pedigree-recording and rely on genomic predictions (e.g., multiplier flocks) is to not include these sheep in NZGE, but make predictions only based on their genotypes using separate analyses.

Dealing with a high number of genotyped animals

So that the evaluation can still be run when the number of genotypes animals exceeds about 150,000, other strategies to estimate the **G** and to invert or solve it may be needed. This is because calculating and inverting of **G** is computationally demanding, where the computations increase about cubically with the number of genotypes. Here we consider some future computing options.

We investigated the algorithm for proven and young animals (APY)(Misztal et al. 2014) (Fragomeni et al. 2015), that gives an approximation to $\mathbf{G}^{-1}$. In general, the results were satisfactory with a core of animals selected randomly (Nilforooshan & Lee 2019). However, the APY was not used in the NZGE, because breeding values from some flocks re-ranked between successive evaluations with only slight changes in the core. It was problematic to resolve this even with a tailored core (results not included). It was concluded that the population being evaluated was too heterogenous for APY to be reliably used. Focus was therefore moved to exact methods rather than approximations.

In exact ssGBLUP (Fernando et al. 2016), the genotyped animals were partitioned into core and noncore animals. The MME used were constructed with only the breeding values of the core animals and breeding values of the non-core animals were estimated as a linear function of the estimates of the breeding values of the core animals. This gives exact solutions for all animals. It is considered that this method should not be any more difficult to implement than APY, however, this was not proven for a large dataset and would require significant changes to the NZGE system.

Another method ssGTBLUP (Mäntysaari et al. 2017), uses a T-matrix and is equivalent to ssGBLUP, but is computationally less demanding. Furthermore, the T-matrix has dimensions $m \times n$, where m is the number of markers and n the number of genotyped animals. This is in contrast to the H-inverse that has dimensions $n \times n$. This becomes important with large numbers of genotypes in the evaluation as the memory usage to store these matrices becomes large. ssGTBLUP is being further investigated as a longer term solution for NZGE.

An alternative strategy is to use brute force for matrix inversion. For example, a **G** of 2,097,152 × 2,097,152 matrix (of size 32TB) could be inverted in 6.4 hours on 2,048 compute nodes (49,152 cores) (Gray, 2016). However, from a practical and economic perspective this approach is not preferred. Another way of overcoming the bottleneck due to large numbers of genotypes is to construct and invert **G** in partitions. This uses mathematical operations for block matrices and recognises that a square matrix can be inverted by partitioning, where the largest block to solve is n divided by the number of partitions. **G** can be constructed and solved for 200,000 genotyped animals partitioned into 4×4 blocks on a modest machine (results not shown).

Hardware and analysis

The NZGE is run on cloud-based computer services and is automated. This is facilitated using docker containers which are orchestrated for weekly evaluations. The use of containers adds portability and ease of automation, where in principle, it can be used on any server or operating system. The use of cloud-based compute infrastructure has allowed immediate and seamlessly change in machine types and number of machines. The NZGE uses about 20 high-performance computing machines (typically 128-512 GB of RAM and 16-48 processors). An evaluation with about 150,000 genotyped animals and about 10 million animals running up to 20 different goal trait groups models will typically take about 20 hours to run, currently, including pre-processing of data.

Conclusion

An overview of the current NZGE system and some of the issues in developing and maintaining this evaluation system, particularly regarding when genotype data is included. The population evaluated in NZGE is heterogeneous and includes different breeds and stabilised crosses of these breeds. The major proportion are Romney-based and genomic prediction for these breeds has been integrated into NZGE. The foundation for NZGE is decades of data recording and models and parameter estimation by many researchers. The implementation of genomic prediction has allowed more utility to be gained from this foundation.

In the future, new and/or significantly more data may need to be evaluated such as data from whole genome sequencing, markers associated with QTL and data from non-seedstock farms and processors (e.g. from image analysis using artificial intelligence). A significant challenge remains to include disparate breeds into a single genetic evaluation with genomic predictions.

While genomic selection is a useful tool for breeders, it is important to be mindful that data recording is still the foundation for genetic improvement which will then allow breeders to best leverage genomic technologies for their selection decisions.

Acknowledgements

We acknowledge the commitment of many individuals and organisations involved in data recording, research and development and for financial investment that forms the basis of NZGE. Meat levy payers through Beef+Lamb New Zealand have supported and funded development and routine NZGE. Supported by the Ministry of Business, Innovation and Employment (NZ) is also acknowledged. We thank Dr Neville Jopson for useful comments related to this paper.

References

Aguilar I, Misztal I, Legarra A, Tsuruta S 2011. Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. Journal of Animal Breeding and Genetics 128: 422-428.

Auvray B, McEwan JC, Newman SA, Lee MA, Dodds KG 2014. Genomic prediction of breeding values in the New Zealand sheep industry using a 50K SNP chip. Journal of Animal Science 92: 4375-4389.

Falconer DS 1996. Introduction to Quantitative Genetics. Pearson Education.

Fernando RL, Cheng H, Garrick DJ 2016. An efficient exact method to obtain GBLUP and single-step GBLUP when the genomic relationship matrix is singular. Genetics Selection Evolution 48: 80.

Fragomeni BO et al. 2015. Hot topic: Use of genomic recursions in single-step genomic best linear unbiased predictor (BLUP) with a large number of genotypes. Journal of Dairy Science 98: 4090-4094.

Garrick DJ, Taylor JF, Fernando RL 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. Genetics Selection Evolution 41: 55.

Gray A 2016. Invertastic: Large-scale Dense Matrix Inversion, EPPC, The University of Edinburgh. June, 2016. https://www.archer.ac.uk/documentation/white-papers/invertastic/invertasticGray.pdf

Gilmore A, Gogel BJ, Cullis BR, Thompson R, Butler D 2009. ASReml User Guide Release 3.0. In: V. I. Ltd (ed.), Hemel Hempstead, United Kingdom.

Harris BL, Johnson DL 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. Journal of Dairy Science 93: 1243-1252.

Henderson CR 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics 31: 423-447.

Henderson CR, Kempthorne O, Searle SR, von Krosigk CM 1959. The Estimation of Environmental and Genetic Trends from Records Subject to Culling. Biometrics 15: 192-218.

Mäntysaari EA, Evans RD, Strandén I 2017. Efficient single-step genomic evaluation for a multibreed beef cattle population having many genotyped animals. Journal of Animal Science 95: 4728-4737.

Meuwissen TH 2009. Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. Genetics Selection Evolution 41: 35.

Misztal I, Legarra A, Aguilar I 2014. Using recursion to compute the inverse of the genomic relationship matrix. Journal of Dairy Science 97: 3943-3952.

Newman SA, Dodds KG, Clark JN, Garrick DJ, McEwan JC 2000. The Sheep Improvement Limited (SIL) genetic engine. Proceedings of the New Zealand Society of Animal Production. 60: 195-197.

Nilforooshan MA, Lee M 2019. The quality of the algorithm for proven and young with various sets of core animals in a multibreed sheep population. Journal of Animal Science 97: 1090-1100.

Pickering NK, Oddy VH, Basarab J, Cammack K, Hayes B, Hegarty RS, Lassen J, McEwan JC, Miller S, Pinares-Patiño CS, de Haas Y 2015. Animal board invited review: genetic possibilities to reduce enteric methane emissions from ruminants. Animal 9: 1431-1440.

Sargolzaei M, Chesnais JP, Schenkel FS 2014. A new approach for efficient genotype imputation using information from relatives. BMC genomics 15: 478-478.

VanRaden PM 2008. Efficient methods to compute genomic predictions. Journal of Dairy Science 91: 4414-4423.