The impact of heterogeneous residual variance across years on estimated breeding values for fly strike in sheep

Elena Dehnavi[1], Andrew A Swan[1], Grant Burbidge[2], Daniel J Brown[1]

1. AGBU, a joint venture of NSW Department of Primary Industries and University of New England, Armidale, NSW, Australia

2. Burbidge Farms, Tarcutta, NSW, Australia

Corresponding author: Elena Dehnavi. Email: edehnavi@une.edu.au

Introduction The accurate genetic evaluation of livestock requires accounting for possible heterogeneity in data to accommodate differences across groups of animals in the levels of residual variance that might exist between years, flocks and locations. Genetic heterogeneity may also exist due to the environments creating different expressions of traits (Mulder et al. 2007). Ignoring the heterogeneity of variances may lead to incorrect ranking and selection of sires and dams, loss of expected genetic gain, and decreased breeding program efficiency (Henderson, 1975). This study was conducted to examine the presence of heterogeneous residual variance in breech strike records as well as the impact of pre-analysis standardisation of data on reducing heterogeneity of variance and accuracy of prediction of genomic estimated breeding values (GEBVs) in the Connemara Merino ram breeding flock.

Material and methods The dataset included breech strike phenotypes (0 or 1) of 1941 lambs born between 2017 and 2021 in the "Connemara" Merino ram breeding flock (Dehnavi et al. 2023). To assess the presence of heterogeneous residual variance, the dataset was divided into two subsets based on contemporary groups (CGs), using a threshold of 0.2 for the incidence rate of breech strike within each CG. Observed data underwent standardisation within CGs using the formula (X-μ)/SD, where X was breech strike phenotype (0 or 1), μ and SD were, respectively, the mean and standard deviation of each CG as defined in Brown et al. 2010). The model incorporated the fixed effects of contemporary groups and the interaction of birth type and rear type while accounting for random effects of direct animal genetic effects. Statistical analyses to test the presence of heterogeneous residual variance involved modelling observed and standardised records of breech strike. A binary trait model with a probit link function was fitted to the observed data, while a linear model was fitted to both observed and standardised data, with both models fitted in ASReml (Gilmour et al. 2015). Genomic information was then integrated using single-step genomic prediction via the Gibbsf90 software (Misztal et al. 2014), and genomic breeding values were estimated. The accuracy of prediction was assessed using a validation procedure in which records of animals born in 2021 were excluded from the analysis. The "partial" GEBVs derived from this subset (300 animals born in 2021) were then compared to the GEBVs obtained from the full analysis to assess accuracy (correlation), dispersion (regression slope) and bias (average difference between partial and full GEBVs). The criteria were computed separately for progeny and their sires.

Results The study revealed the presence of heterogeneity in residual variance across years of birth, with a residual variance of 0.20±0.01 and 0.06±0.01 for high-incidence (a mean greater than 0.2) and low-incidence breech strike (a mean less than 0.2) subsets, respectively, when running a linear model on observed data. After standardising for CGs, these variances shifted to 0.94±0.05 and 0.91±0.06, indicating effective mitigation of heterogeneity. However, no significant evidence of heterogeneity was found in a binomial model, with values of 0.93±0.05 and 0.88±0.08 in the high and low-incidence subsets, respectively.

Results from the validation analysis showed a significant increase (P < 0.05) in the prediction accuracy when employing a linear model on standardised data compared to running a binomial model on observed data. This was evidenced by an increased Pearson correlation for both progeny and sires. Minor, non-significant differences in dispersion and bias were noted.
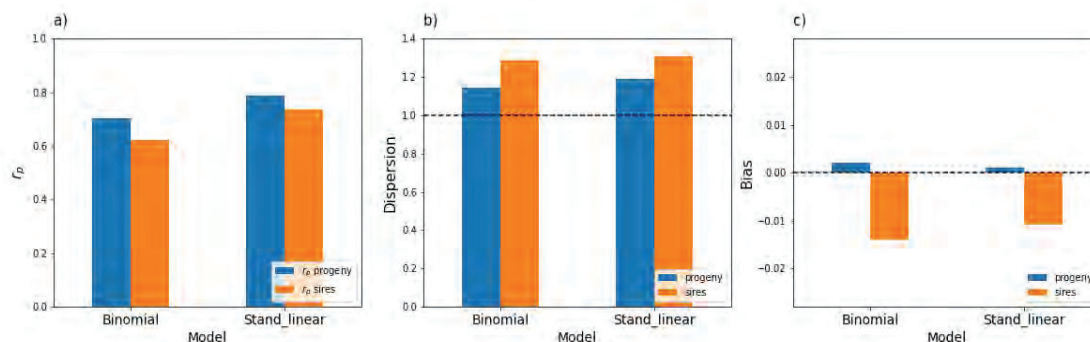

Figure 1. Comparison of the accuracy and bias of predictions obtained by applying the linear model to standardised data (Stand_linear) and the binomial model to observed data. This illustrates accuracy (r_p) (a), dispersion (b), and bias (c). Results are presented for progeny (blue bars) and their sires (orange bars).

Conclusion and implications There was evidence of heterogenous variance using a linear model on observed data, which could be effectively addressed by standardising records within contemporary groups. There was no significant evidence of heterogeneity when using a binomial model, and the GEBVs were generally comparable to those obtained from the linear model with standardised data. These findings are essential for implementation in genetic evaluation systems because linear models are more efficient computationally and more flexible in fitting complex models, including multi-trait models.

References

Brown D. J., Swan A.A. and Gill J.S. 2010. Within- and across-flock genetic relationships for breech flystrike resistance indicator traits. Animal Production Science. 50: 1060-1068.

Dehnavi E., Swan A.A., Ramsay A.M.M., Bird-Gardiner T.L. and Brown D.J. 2023. Preliminary genetic parameters for flystrike and its association with production traits in Australian Merino sheep. In Proceedings of the 15th Conference of the Association for the Advancement of Animal Breeding and Genetics. Perth, WA, 25: 354-357.

Gilmour A.R., Gogel B.J., Cullis B.R., Welham S.J. and Thompson R. 2015. ASReml user guide release 4.2 structural specifications. VSN International Ltd: Hemel Hempstead, HP1 1ES, UK.

Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics. 31:423.

Misztal I., Tsuruta S., Lourenco D.A.L., Aguilar I., Legarra A. and Vitezica Z. 2014. Manual for BLUPF90 family of programs.

Mulder H.A., Bijma P. and Hill W.G. 2007. Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. Genetics. 175:1895–1910.