Smith, J., Sneddon, N., Jopson, N., Martín, N., López-Villalobos, N., van der Linden, D., (2026) Investigating the impact of genetic linkage between sheep flocks on the accuracy and outcomes of genetic evaluations, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 24382: https://doi.org/10.31274/wcgalp.24382

Investigating the impact of genetic linkage between sheep flocks on the accuracy and outcomes of genetic evaluations

J.L. Smith[1], N.W. Sneddon[1*], N.B. Jopson[2], N.P. Martín[1], N. López-Villalobos[1], and D. van der Linden[3]

[1]Massey University, Palmerston North, New Zealand; [2] AbacusBio Limited, Dunedin, New Zealand; [3] Beef + Lamb New Zealand Genetics, Dunedin, New Zealand; [*]Corresponding author: N.W.Sneddon@massey.ac.nz

Introduction

Genetic gain in the New Zealand sheep industry relies on the accurate estimation of breeding values (BVs) for different traits and the ability to compare them reliably across flocks. The effectiveness of across-flock BV comparison hinges on genetic connectedness, defined as the degree of genetic linkage between animals in different flocks (Kuehn *et al*., 2007). Sire sharing is the primary mechanism used to establish this linkage, enabling animals raised in different environments to be evaluated on a common genetic scale (Brown *et al.*, 2007). Despite the importance of linkages, there are few evidence-based guidelines specifying: the number of linkages required, the effectiveness of historical links, and thresholds to support accurate comparisons of BVs across flocks. In New Zealand, the nProve evaluation applies a relatively conservative connectedness criterion, assessing linkage based on the number of shared progeny from link sires within a rolling three-year window (Newman, 2003). Although this approach ensures robust benchmarking, it may restrict the number of connected flocks. This could be of advantage for traits that are difficult to measure. For example, facial eczema tolerance (FET) has limited phenotypic recording despite being a serious health and production issue in New Zealand. Facial eczema is a fungal disease caused by ingestion of the toxin sporidesmin (Morris *et al*., 2004). Phenotyping for FET is expensive and difficult to measure due to the controlled nature of the challenge and risk to animals, limiting recording to a subset of animals and number of participating flocks (B+LNZ Genetics, 2025).

This study evaluates how inclusion criteria and data composition influences connectedness statistics. By identifying thresholds for connectedness that support reliable across-flock comparisons, the aim was to inform progeny test designs and evaluate whether current connectedness criteria could be relaxed while still maintaining valid genetic benchmarking.

Materials & Methods

Data for this study were obtained from the nProve database, which included FET, pedigree, and identification records for all animals in the system. Data processing and filtering were carried out in SAS 9.4 using the standard nProve preprocessing programme to generate pedigree, unique identification, and trait files for connectedness assessment.

Connectedness was first evaluated using the current nProve methodology, which computes across-flock linkage based on the number of progeny from link sires shared between each pair of flocks within a three-year window appropriate to each trait. This method constructed a flock-by-flock connectedness matrix, in which linkage strength reflects the total number of progeny from link sires used across both flocks. This matrix was transformed into a distance matrix and visualised using hierarchical clustering, producing a dendrogram that identifies groups of connected and unconnected flocks. For this purpose, any flocks with connectedness value <0 are considered unconnected. The second approach used the Genetic Connectedness Analysis package in R (Yu and Morota, 2021). A pruned pedigree was generated for animals contributing FET data in the last three years, and a corresponding relationship matrix was constructed.

Three different analyses were conducted to assess how changes to inclusion criteria and data composition influence observed connectedness. First, the assessed data window was varied (from 1 year to 6 years). Second, highly connected flocks were removed from analysis, one by one, to evaluate the stability of across-flock linkage and the dependence of the system on key connecting flocks (flocks F1, F6 and F18, with connectedness values >2.2, >2.4 and >1.6, respectively). Third, the three most prolific sires were identified and removed one by one to assess how the loss of major link sires altered flock connectedness. For each modified scenario, connectedness matrices and dendrograms were generated using the nProve procedure and compared with the baseline connectedness structure.

The Genetic Connectedness Analysis (GCA) package in R was also used to quantify across-flock connectedness under the different scenarios tested using Best Linear Unbiased Prediction (BLUP) equations and relationship matrices (Yu and Morota, 2021). To assess the impact of the number of years of data on the coefficient of determination (CD) and prediction error variance of differences (PEVD) for the accuracy of genetic evaluations across flocks, the GCA connectedness analysis was re-run in RStudio using progressively larger subsets of the FET dataset. The model fitted included the fixed effects of flock and sex, σ_a^2= 0.3321 and σ_e^2= 0.4779. The range of birth years included was incrementally expanded from two years of phenotypic data (earliest birth year included = 2023) up to ten years (earliest birth year included = 2015). Across-flock evaluation accuracy was quantified using PEVD and CD. Pairwise PEVD values between flocks were first calculated using the GCA framework and then summarised as the overall average PEVD ($PEVD_{GrpAve}$) across all flock comparisons. The corresponding overall average CD (CD_{GrpAve}) was also calculated, as the variance of true differences between flocks computed from σ_a^2 and the mean of pairwise relationships within and between flocks derived from the pedigree matrix. For each scenario, overall $PEVD_{GrpAve}$, CD_{GrpAve}, the number of phenotypic records, and the number of animals retained in the pruned pedigree were recorded.

Results & Discussion

Genetic connectedness among flocks, as assessed using the nProve method, is shown in Figure 1. The number and percentage of connected flocks are presented in Table 1. Clusters with shorter branch lengths indicate stronger genetic linkage between flocks. The scale ranges from 3 to -2, with positive values meeting connectedness criteria, and negative values do not. Based on this threshold, 24 of 49 flocks were classified as connected for the base scenario, forming a cluster above flock F25, with the remaining 25 flocks unconnected.

Table 1. Number and percentage of flocks identified as connected (flocks that met the connectedness criteria, connectedness >0) or unconnected, as inferred from the dendrograms generated under each scenario in the nProve connectedness analysis for facial eczema (FET).

Scenario	Total flocks (n)	Connected flocks (n)	Unconnected flocks (n)	% Connected
Base (3y window)	49	24	25	49.0
1y window	46	10	36	21.7
2y window	48	12	36	25.0
4y window	51	34	17	66.7
5y window	55	35	20	63.6
6y window	58	41	17	70.7
Base - Flock 1	48	20	28	41.7
Base - Flock 6	48	22	26	45.8
Base - Flock 18	48	23	25	47.9
Base - Prolific Sires	49	24	25	49.0

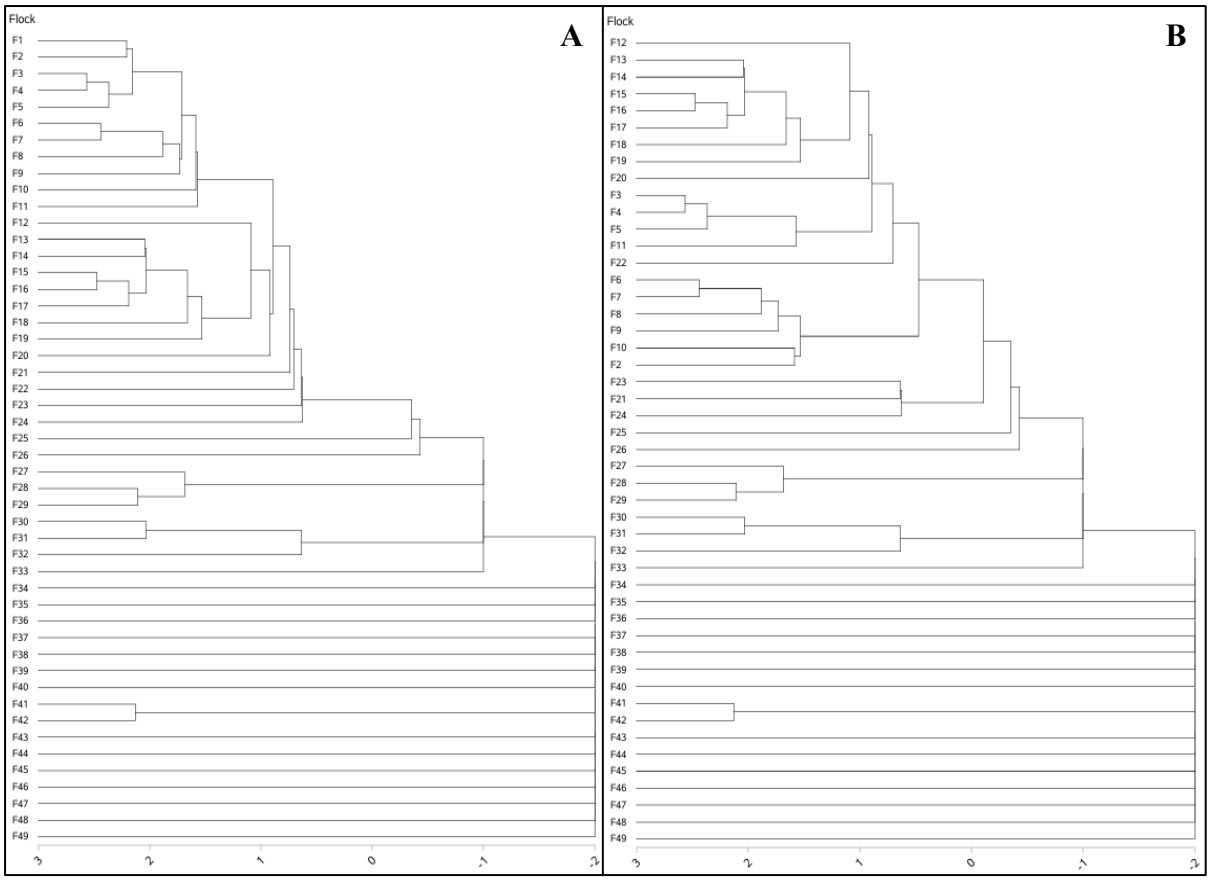

Figure 1. Dendrograms illustrating genetic connectedness among flocks for facial eczema. Clusters with shorter branch lengths indicate stronger genetic linkage between flocks using the nProve method (A) Base scenario (3-year window, no flocks or sires removed); (B) Flock F1 removed but same inclusion criteria as the base scenario.

Effect of data-window on connectedness. The current recommendation is to restrict across-flock linkage assessments to a 3 to 5 year window (Newman, 2003); however, an expansion of this range has not previosusly been evaluated in relation to connectedness or across-flock accuracy. The number of connected flocks increased as the data window was extended (Table 1). Reducing the window from the current 3-year standard data window to a 2- and 1-year window resulted in a decline in overall connectedness. This reduction was reflected in the proportion of connected flocks, which fell from 49.0% to 25.0% and 21.7% for 2- and 1-year windows, respectively. In contrast, extending the window to 4-, 5-, and 6-years increased connectedness, with 70.7% of flocks connected when a 6-year window was used.

Flock and sire removal. As this has not been evaluated previously, sequential removal of highly connected flocks altered the dendrogram structure and changed the connectedness status of multiple flocks, with some immediately losing any connection to other flocks within the evaluation as shown in Figure 1. Removal of flock F1 reduced the proportion of connected flocks from 49.0% to 41.7%, while removal of flock F6 reduced connectedness from 49.0% to 45.8% (Table 1). In contrast, removing flock F18 had minimal influence on dendrogram structure or overall connectivity. The removal of the three most prolific sires in the FET dataset also produced negligible changes, with no flock shifting in or out of the connected status, indicating that no individual sire maintained across-flock connectivity in this evaluation.

Effect of data-window on PEVD and CD. Across-flock evaluation accuracy for FET measured using PEVD and CD showed little change as the data window increased (Figure 2). Using a 3-year dataset consistent with the data window currently applied in nProve for assessing connectedness (earliest birth year = 2022), PEVD was 0.564 and CD was 0.286. Including older data increased the number of phenotypic records from 4,931 to 25,890, and the number of animals in the pedigree from 12,076 to 60,081 with 10 years of data. CD increased smoothly by 0.112 over the 10-year window (to 0.371), indicating improved across-flock accuracy, whereas PEVD fluctuated within a narrow range (0.504 – 0.587) across the range of data windows evaluated. Both metrics showed evidence of plateauing after approximately 5 to 6 years of data (earliest birth year = 2019-2020).

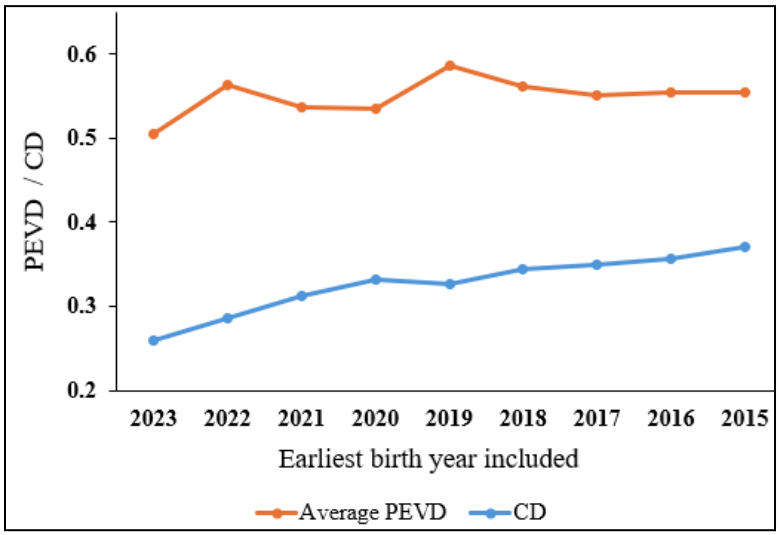

Figure 2. Effect of data-window length on across-flock evaluation accuracy for facial eczema (FET), measured using the average prediction error variance of differences ($PEVD_{GrpAve}$) and the coefficient of determination (CD_{GrpAve}). The x-axis shows the earliest birth year included in the evaluation.

Conclusions

Extending the assessed data window beyond the current 3-year standard increased the number of connected flocks by capturing additional historic genetic links. However, it was unclear whether the inclusion of these historical connections would improve the validity of across-flock BV comparisons. Analysis of evaluation accuracy using the GCA method demonstrated that extending the data window did not improve PEVD; however, CD increased, indicating improved reliability of across-flock comparisons. These findings suggest that for FET, extending the nProve evaluation to a 5-year window could enable greater flock participation and provide a practical balance between improving CD, maintaining PEVD, and keeping the evaluation manageable in terms of computational time and effort. Whether these findings extend to traits with higher levels of phenotypic recording remains to be evaluated.

References

B+LNZ Genetics. (2025). Facial Eczema Tolerance: Technical Note. https://www.sil.nz/files/174674064762.pdf

Brown D., Huisman A., Swan A., Graser H., Woolaston R., Ball, A., Atkins, K., and Banks, R. (2007) Proc. Assoc. Advanc. Animal Breed. Genet., 17: 187. https://www.aaabg.org/livestocklibrary/2007/brown187.pdf

Kuehn, L.A., Lewis, R.M., and Notter, D.R. (2007) Genet. Sel. Evol. 39:225 doi.org/10.1186/1297-9686-39-3-225

Morris, C., Towers, N., Hohenboken, W., Maqbool, N., Smith, B., and Phua, S. (2004) New Zealand Veterinary Journal 52: 205-215. doi.org/10.1080/00480169.2004.36431

Newman, S.-A. (2003). Proceedings of the New Zealand Society of Animal Production, 63: 194-196.

Yu, H., and Morota, G. (2021) BMC Genomics, 22: 119. doi.org/10.1186/s12864-021-07414-7