Proc. Assoc. Advmt. Anim. Breed. Genet. 26: 471-474

DATA QUALITY BY GEOGRAPHICAL LOCATION FOR AUSTRLIAN AND NEW ZEALAND MERINOSELECT FLOCKS

C.J. Bunter, P.E. Bradley, M.J. Hodge and K.L. Rummery

Meat and Livestock Australia, University of New England, Armidale, NSW, 2351 Australia

SUMMARY

The rate of genetic gain can be maximised by making selection decisions using accurate breeding values, underpinned by good quality data. This paper identifies and examines variation in data quality metrics in MERINOSELECT flocks across different geographic regions in Australia and New Zealand. Data quality metrics with significant variation across regions included pedigree completeness, variation in age and birth type of animals within a cohort, recording of fleece traits, and across flock linkage for reproduction traits. Improving such metrics via targeted adoption and extension messaging may lead to improved data quality, which may ultimately improve the rate of genetic gain resulting in more productive and profitable sheep.

INTRODUCTION

Maximising genetic gain in Merino flocks requires selection decisions to be informed by accurate breeding values, underpinned by high-quality data (Guy and Brown 2021). The Data Quality Score (DQS) report, was developed to help Sheep Genetics breeders identify key data quality metrics for improvement, and as a result, drive increased rates of genetic gain (Brown *et al.* 2022). Since the release of the DQS in 2022, anecdotal evidence suggests that there are regional differences in data quality (Sheep Genetics 2022).

The DQS presents as a score out of 100 for each flock and a star rating out of 5, considering the quantity and quality of data provided by the breeder over a 5-year rolling average. These metrics are weighted based on their relative importance (Guy and Brown 2021). This paper investigates the difference in individual DQS metrics for MERINOSELECT flocks between geographic regions, aiming to inform more targeted extensions and adoption campaigns.

MATERIALS AND METHODS

The 7th October 2024 MERINOSELECT run DQS metrics were used for all MERINOSELECT flocks. A flock's data was removed from the dataset if it did not have more than 3 years of current data, was non-active (such as research or legacy flocks), was from a non-Merino breed (such as Polwarth), or located outside of Australia and New Zealand. A total of 267 flocks were included in the final dataset.

DQS metrics were extracted and expressed as percentages including: i) proportion of full pedigree (FullPed); ii) average pedigree (AvgPed); iii) average recording of measurements of weight (RecWT), reproduction (RecRepro), wool (RecFW), carcase (RecCA) traits as percentage of the drop recorded; iv) proportion of syndicate pedigree (SynPed); v) proportion of age variation within a contemporary group (Ages); vi) proportion of birth type variation within a contemporary group (BT); vii) effective progeny as the percentage of progeny from a given sire relative to all progeny within a group (Brown *et al.* 2001) (Eff); and, viii) pedigree linkage to the analysis for weight (LinkWT), reproduction (LinkRepro), wool (LinkFW) and carcase traits (LinkCA) as a measure of the proportion of animals with a half sibling in another flock recorded for that trait group.

Flocks were assigned a region based on postcode as defined by the Australian and New Zealand governments (NSW Government 2025; NZ Immigration 2025; RDA 2025). Any regions with fewer than 3 flocks were combined into their relative states' 'other' category, allowing for better comparisons (Table 1). A series of one-way ANOVA tests were conducted to identify significant variation for data quantity and quality metrics between regions.

Table 1. Count of flocks per region, with region number as referenced in Figure 1

	Region	Number of flocks		Region	Number of flocks
1	Barwon south west, VIC	22	13	Limestone coast, SA	13
2	Canterbury, NZ	15	14	Loddon Mallee, VIC	8
3	Capital Region, ACT	15	15	Marlborough, NZ	4
4	Central West, NSW	14	16	Murraylands and Riverland, SA	8
5	Queensland	3	17	New England and North West, NSW	18
6	Eyre Peninsula, SA	11	18	Otago, NZ	9
7	Far West and Orana, NSW	13	19	Riverina, NSW	10
8	Murray, NSW	10	20	Tasmania	9
9	Goldfields Esperance, WA	6	21	Wheatbelt, WA	27
10	Great Southern, WA	20	22	York and Mid north, WA	8
11	Hume, VIC	11	23	Western Australia other	4
12	Kangaroo Island, SA	5	24	Victoria other	4

RESULTS AND DISCUSSION

The DQS identified to have significant variation (P<0.05) between regions by one-way ANOVA are displayed in Figure 1. Greater levels of variation between regions were identified compared to variation within regions for metrics FullPed, AvgPed, SynPed, Ages, BT, RecFW and LinkRepro.

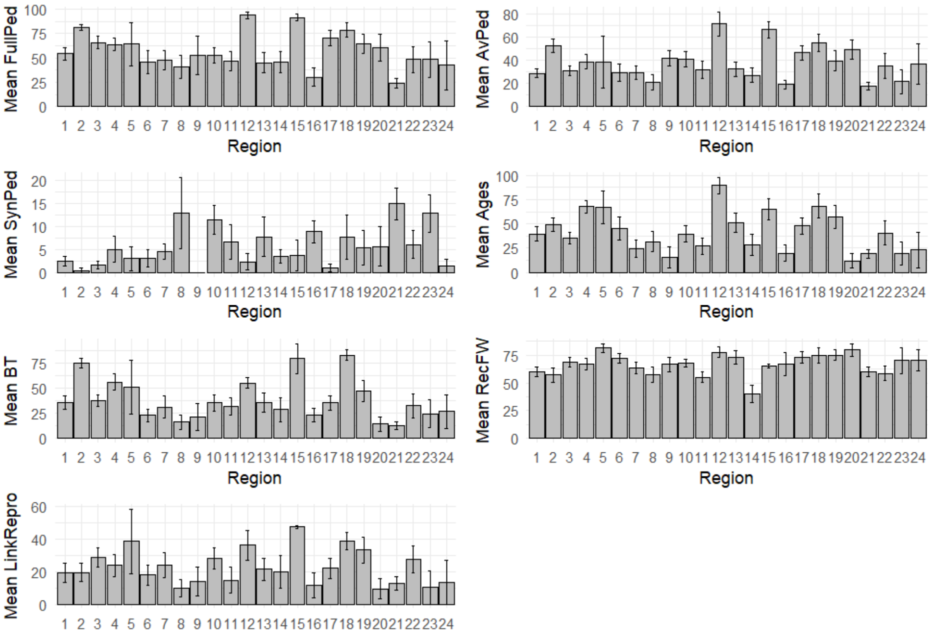

Figure 1. Average (box) and range (± standard error bar) for each data quality metric by region

Pedigree. Increased recording and depth of sire, dam and back pedigree provides more information for the calculation of breeding values, therefore increasing their accuracy. FullPed refers to the proportion of animals where both sire and dam are known. The average FullPed across regions was 56.6%, up on a previous study where Merino flocks averaged 43.6% in 2022 (Brown *et al.* 2022). Five regions averaged over 70% FullPed and two regions (Murraylands and Riverland, and Wheatbelt) averaged under 35% (Figure 1). AvgPed refers to the average depth of pedigree known. The average AvgPed was 37.4%, with the same 2 regions as FullPed averaging under 20% AvgPed (Figure 1). SynPed refers to the proportion of animals with syndicate sire pedigree, rather than an individual sire. The average SynPed was 5.5%, but 4 regions averaged greater than 10%: Murray, Great Southern, Wheatbelt and WA other (Figure 1).

Merino flocks historically have lower levels of pedigree recording compared to other breeds, but the increase in FullPed identified in this study could be attributed to the exponential uptake of genomic technology informing pedigree in recent years (Sheep Genetics 2024). While there has been an increase in pedigree recording, the previously mentioned regions still have low pedigree recording. Pedigree recording is integral to genetic evaluation and as a result, pedigree metrics are highly weighted in the DQS. This means that flocks that accurately provide full pedigree generally have a higher DQS. Pedigree is one of the most important components of genetic evaluations and as such, the Merino sheep industry could see some improvements by targeting resources towards improving pedigree recording and recording methods in Murraylands and Riverland, Wheatbelt, Murray, Great Southern, and WA other.

Fixed effects. Fixed effects are known non-genetic factors that have an impact on phenotypic expression and are therefore adjusted for or fitted in the model of the MERINOSELECT genetic evaluation. The accurate recording of fixed effects results in a more accurate breeding value calculation.

Ages refers to the proportion of age variation within a contemporary group, where a higher percentage indicates more variation in date of birth recorded. The average Ages was 39.4%, with 7 regions averaging under 25%, and 17 regions under 50% (Figure 1). This average is consistent with Brown *et al.* (2022), who found Merinos had significantly less age variation compared to LAMBPLAN flocks. Limited age information remains a difficulty, with little improvement from previous studies.

BT refers to the proportion of birth type variation within a contemporary group, where a higher percentage indicates more variation in birth type recorded. The average BT was 36.1% with 7 regions under 25% (Figure 1). BT may be supplied from pregnancy scan result, genomic testing identifying number of lambs per dam and count of lambs at birth. Evidence from the Sheep Genetics database suggests that in MERINOSELECT, BT is most commonly inferred from pregnancy scanning result. The differences in recording methods may explain some of the variation identified between regions. Targeting the nine regions that were below 25% for either Ages or BT with education specific to fixed effect recording would likely see an increase in the recording of Ages and BT, resulting in more correct breeding values and therefore a higher rate of genetic gain.

Phenotypic recording. RecFW refers to the proportion of animals recorded for fleece traits. The average RecFW was 66.8%, with only one region below 50%: Loddon Mallee (Figure 1). The recording of fleece traits is moderately weighted in the DQS, as phenotypic recording at a flock level provides information to the calculation of breeding values, increasing their accuracy, and it also helps inform the reference population for these traits. Phenotypic recording can vary across flocks based on the importance of that trait to each flock's breeding objective.

Linkage. LinkRepro refers to the proportion of animals that have a half sibling recorded for reproduction traits in another flock. The average LinkRepro was 22.6%, with Murray and Tasmania recording less than 10% LinkRepro, compared to Marlborough NZ averaging 47.5% (Figure 1). Linkage is integral to across flock genetic evaluation, allowing for the accurate comparison of animals in different environments. The variation in LinkRepro could be explained by differences in breeding objectives and therefore recording of reproduction traits. The current reproduction analysis (Weaning Rate analysis) used in MERINOSELECT was introduced in 2022 (Sheep Genetics 2022) and as the number of generations recorded for the Weaning Rate analysis increases (rather than the previously used NLB/NLW reproduction analysis), LinkRepro may increase as a result.

Targeting extension messaging. This analysis enables more targeted genetic extension activities, focusing on regions with lower performance in key data quality metrics, linked to genetic gain. By identifying which regions perform poorly in certain metrics, Sheep Genetics can prioritise targeting regions with content specific to their current weaknesses, increasing genetic gain. This allows for best use of time and resources, resulting in the largest impact. Additional work will need to be done to replicate this investigation with LAMBPLAN flocks, better identifying the variation between regions in data quality metrics that influence genetic gain across all Sheep Genetics evaluations.

CONCLUSION

Significant variation between regions was found for DQS metrics FullPed, AvgPed, SynPed, Ages, BT, RecFW and LinkRepro, suggesting there are region specific methods of data recording. The results from this paper may be used for region specific extension and education programs to improve data quality. This may produce a subsequent improvement in genetic gain for Australian and New Zealand Merino flocks. While genetic gain is not only impacted by data quality, this investigation provides Sheep Genetics with tangible outcomes to achieve better data quality in MERINOSELECT and therefore potentially increase future genetic gain.

REFERENCES

Brown D.J., Tier B. and Banks R. (2001) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **14**: 441.

Brown D.J., McCrabb E.J., Bradley P.E., Rose I.J., Banks R.G. and Guy S.Z.Y. (2022). *Proc World Cong. Genet. Appl.Livest. Prod.* **12**: 743

Guy S.Z.Y and Brown D.J. (2021) *Proc. Assoc. Advmt. Anim. Breed. Genet.* **24**:143

New South Wales Government. (2025). *Smart and skilled regions.* https://www.nsw.gov.au/education-and-training/resources/smart-and-skilled-regions

New Zealand Immigration. (2025). *Regions and cities*. http://live-work.immigration.govt.nz/choose-new-zealand/regions-cities

Regional Development Australia. (2025). Regional Development Australia. https://www.rda.gov.au/

Sheep Genetics. (May, 2022). “Analysis Enhancements 2022”. https://www.sheepgenetics.org.au/globalassets/sheep-genetics/updates/analysis-enhancements/analysis-enhancements-2022.pdf

Sheep Genetics. (2024). “Sheep Genetics Annual Outcomes Report”. https://www.sheepgenetics.org.au/globalassets/sheep-genetics/about-us/annual-report-2023-24.pdf