Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Evaluation of Daily Milk Production and Liveweight Records in New Zealand Dairy Cattle: Implications for Genetic Improvement and Breeding Strategies.

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Science

In

Animal Science: Genetics and Breeding



at Massey University, Manawatu, New Zealand.

Korbyn Taylor Patricia Garland

2026

ABSTRACT

Understanding the relationships between live weight (LWT) and milk production traits is important for improving biological and economic efficiency, defined as the ability of cows to convert feed into saleable milk solids (MS) while maintaining body condition and reproductive performance, in pasture-based dairy systems. This study examined how LWT, milk yield (MY), and milk composition interact across lactation, with emphasis on the biological constraints and the implications relevant to the nature of the New Zealand dairy industry. Repeated test-day records from a commercial herd were analysed to capture within-cow variation across a single lactation season. Mixed-effect models were used to account for temporal dependence, recognising that repeated measurements collected from the same cow over time are not independent and are more strongly related than those recorded further apart. Phenotypes adjusted for LWT, age at calving, lactation stage, and within-cow temporal effects were used to investigate intrinsic relationships among traits. Live weight was positively associated with MY and MS, indicating greater production capacity in larger cows; however, this relationship was non-linear, with diminishing returns at greater body weights. A consistent negative relationship between MY and milk component concentrations was observed after adjustment, confirming the persistence of a dilution effect independent of body size and lactation stage. Temporal patterns showed that milk production followed the expected non-linear lactation curve, characterised by an early increase followed by a gradual decline as lactation progressed. Age effects reflected growth constraints in younger cows, while breed comparisons highlighted differences between volume-driven and concentration-driven production strategies. These findings support genetic selection strategies that balance production, body condition, body size, and efficiency. In pasture-based systems where payment is based on milksolids, prioritising efficient conversion of feed into fat and protein is more suitable than optimising milk volume alone.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr Nick Sneddon for his invaluable guidance, constructive feedback, and continuous support throughout the course of this research. His expertise in animal breeding and genetics, encouragement, and patience were instrumental in shaping both the direction and quality of this thesis. I am also especially grateful for his sense of humour throughout my Massey education journey, as he consistently kept spirits high and reminded me to find joy even when challenges felt overwhelming.

I am deeply thankful to Dr Penny Back for her insightfulness, emotional support, and academic guidance. Her contributions significantly strengthened this work. I am particularly appreciative of the financial support she provided through research assistant opportunities during semesters and summer breaks, which not only supported me practically but also extended my knowledge and broadened my experience within animal science.

I would like to acknowledge the Animal Science Department at Massey University for providing the academic environment and resources necessary to undertake this research. I would also like to extend my gratitude to the C Alma Baker Scholarship committee for granting me financial aid during my studies.

Finally, I want to give a special and heartfelt thank you to my mum, Amanda Garland, who put her entire being into raising us kids to the best of her ability and providing us with opportunities she never had. Thank you for being my rock and my best friend.

To my wider family and friends, thank you for your unwavering support, understanding, and motivation throughout this journey. To my employers Brooke and Blair Argyle at Beaver and Bear, thank you for taking me under your wings and employing me for the last 2 ½ years, treating me like family and supporting me throughout my degree. Your support and friendship throughout this time have been invaluable and taught me so much, not to mention the well-appreciated financial support during my studies. Your belief in me provided the foundation that made this achievement possible.

List of Contents

ABSTRACT ..... **i**
ACKNOWLEDGEMENTS ..... **iii**
List of Contents ..... **v**
List of Tables ..... **vii**
List of Figures ..... **viii**
List of Abbreviations ..... **ix**
CHAPTER 1 ..... **1**
1.1 Introduction *3*
CHAPTER 2 ..... **5**
Genes associated with Live Weight, Body Condition Scores and Milk traits in Dairy Cattle and the implications they have on Production: A Literature Review **5**
2.1 Introduction *7*
2.2 Importance of Genome-Wide Studies *8*
2.2.1 Imputation 9
2.2.2 Population Stratification and Cryptic/Family Relatedness 10
2.2.3 Epistasis, Epigenetic and Environmental factors 12
2.3 Genetic Architecture of Live weight and Body Condition in Dairy Cattle *13*
2.3.1 Heritability 13
2.3.2 Quantitative Trait Loci 13
2.3.3 Candidate genes 14
2.3.4 Polygenic nature and gene networks 14
2.3.5 Influence of environment on genetic expression 15
2.4 Genetic Architecture of Milk Traits in Dairy Cattle *15*
2.4.1 Milk Yield 15
2.4.2 Milk Composition 17
2.5 Cross-Trait Relationships *18*
2.6 Current Systems *21*
2.7 Breeding Objectives *22*
2.8 Implications for Breeding Programs *23*
CHAPTER 3 ..... **25**
Methods and Data **25**
3.1 Data *27*
3.1.1 Farm System and Scale 27
3.1.2 Herd Characteristics 27
3.1.3 Data Recording and Monitoring Systems 27
3.1.4 Data Format 28
3.2 Methods *29*
3.2.1 Compiling Data 29
3.2.2 Analysing Data 29
3.3 Genetic Analysis using ASReml *32*
3.3.1 Data Preparation and Pedigree Structure 33

3.3.2 Model Specification ... 33
3.3.3 Variance Components and Genetic Parameters ... 34
CHAPTER 4 ... **35**
Results ... **35**
Dataset Characteristics ... *36*
4.1 Descriptive Temporal Trends and Lactation Patterns ... *36*
4.2 Phenotypic Associations Between Live weight and Milk Production Traits ... *41*
4.3 Linear Mixed Model Results ... *42*
4.5 Adjusted Phenotypes and Trait Relationships ... *44*
4.6 Regression Relationships between Adjusted Phenotypes ... *46*
4.7 Fixed-Effects General Linear Model Results ... *47*
4.8 ASReml Pedigree Results ... *49*
CHAPTER 5 ... **51**
Discussion ... **51**
5.1 Overview of key findings ... *53*
5.2 Relationship between Live weight, Milk Volume and Milksolids ... *53*
5.3 Lactation Stage and Within-Cow Variation ... *55*
5.4 Adjusted Trait Relationships and Component Dilution ... *56*
5.5 Age and Breed Effects on Production and Size ... *57*
5.6 Heritability and Repeatability ... *58*
5.7 Methodological Considerations and Model Performance ... *61*
5.8 Implication and Practical Applications of Findings ... *62*
5.9 Limitations of the Study ... *64*
5.10 Conclusion ... *65*
References ... *67*

List of Tables

Table 1. 1: National herd test production averages for New Zealand dairy cattle during the 2023/2024 and 2024/2025 production seasons. 4

Table 3. 1: Representative sample of cleaned production records used in the analysis. 28

Table 4. 1: Summary statistics of the final analytical dataset per measured trait. 36

Table 4. 2: Phenotypic regression results for associations between live weight and milk production traits. 41

Table 4. 3: Fixed effect estimates (± SE) and autoregressive correlations from linear mixed models for milk production traits. 43

Table 4. 4: Pearson correlations among adjusted milk production phenotypes. 44

Table 4. 5: Least-squares means (± SE) for production traits by age-at-calving group, adjusted for DIM, herd testing date, and breed. 48

Table 4. 6: Least-squares means (± SE) for production traits by breed (Holstein-Friesian (HF), Jersey (J), and Holstein-Friesian Jersey crossbred (HFxJ), adjusted for DIM, herd testing date, and age-at-calving group. 48

Table 4. 7: Variance components (additive (σ^2A), permanent environment (σ^2PE), residual error (σ^2E), phenotypic (σ^2P)) and genetic parameters (heritability (h2), repeatability (t)) for milk production traits. 50

List of Figures

Figure 4. 1: Herd level mean milk production traits and live weight across successive herd test dates. Points represent herd means for each test day, illustrating temporal patterns in milk yield (a), milksolids (b), milk fat percentage (c) and milk protein percentage (d), and live weight (e) across the observation period. 37

Figure 4. 2: Herd-level temporal trends in milk production traits and live weight across days in milk (DIM). Milk volume vs DIM (a), milk fat percentage vs DIM (b), milk protein percentage vs DIM (c), milksolids vs DIM (d), and live weight vs DIM (e). Values represent daily herd means with smoothed curves illustrating non-linear lactation patterns for milk yield, milksolids, milk fat and protein percentage..... 38

Figure 4. 3: Distribution of within-cow changes in milk yield between successive herd test days. Delta_vol represents differences in daily milk yield (kg/day) between consecutive test dates for the same cow, illustrating short-term temporal variability in milk production..... 39

Figure 4. 4: Daily milk yield across days in milk for individual cows. 40

Figure 4. 5: Phenotypic association between; live weight and milk protein percentage (a), live weight and milk fat percentage (b), live weight and milk volume (c), live weight and milksolids (d). Each point represents a test-day observation, with the fitted linear regression line shown. The coefficient of determination (R^2) is reported on the plot. 41

Figure 4. 6: Histogram of residuals from the linear mixed model for milk volume. Residuals are centred near zero and exhibit moderate right skewness with slightly heavier tails than expected under a normal distribution. 45

Figure 4. 7: Relationship between adjusted milk volume and adjusted milk solids. 45

Figure 4. 8: Scatter plot matrix of adjusted milk production traits (milk volume, fat percentage, protein percentage, and milk solids production). 46

Figure 4. 9: Relationship between adjusted milk volume and adjusted milk protein percentage following adjustment for fixed effects. The fitted regression line illustrates the weak but significant negative association between the two traits. 47

Figure 4. 10: Distribution of (a) milk volume and (b) live weight across age-at-calving groups (2–3, 4–5, 6–7, and ≥8 years). 49

List of Abbreviations

1000 BGP	The 1000 Bull Genomes Project.
AR(1)	First-order autoregressive covariance structure.
BCS	Body Condition Score, a measure of an animal's fat cover and energy reserves.
BTA	Bos taurus autosome, which refers to specific cattle chromosomes.
BW	Breeding Worth, the national bioeconomic selection index used in New Zealand to rank cows based on expected profit.
CR42	Calving rate within 42 days, a key measure of reproductive performance in seasonal systems.
DAVID	Database for Annotation, Visualisation and Integrated Discovery, a tool used for functional annotation.
dEBVs	De-regressed estimated breeding values, used to strengthen statistical precision in genetic studies.
DIM	Days in milk, referring to the stage or duration of a cow's lactation.
DMI	Dry matter intake, the amount of feed an animal consumes.
ECM	Energy-corrected milk, a measure that standardises milk yield based on energy content.
eQTL	Expression quantitative trait loci, which are genetic variants that regulate gene expression.
GC	Genomic control, a statistical method used to correct for population stratification.
GH	Growth Hormone.
GHR	Growth Hormone Receptor.
GnRH	Gonadotropin-releasing hormone, involved in hormonal signalling pathways related to milk yield.
GO	Gene Ontology, used in enrichment analyses to translate genetic variants into biological meaning.
GWAS	Genome-wide association studies, which identify genomic regions associated with trait variation.
HF	Holstein-Friesian
HFxJ	Holstein-Friesian x Jersey Crossbred
J	Jersey
IGF1	Insulin-like growth factor 1, a candidate gene and hormone linked to growth and reproduction.
KEGG	Kyoto Encyclopaedia of Genes and Genomes, used for biological pathway enrichment analyses.
LIC	Livestock Improvement Corporation (referenced regarding national dairy statistics).
LMMs	Linear Mixed Models, used to control for genetic relationships in GWAS.
LOCO	Leave-one-chromosome-out, a statistical approach used to prevent overestimation of SNP effects and reduce false positives.
LOESS	Locally estimated scatterplot smoothing.
LWT	Liveweight, a measure of an animal's size and mature growth.
MIR	Mid-infrared spectroscopy, used for cost-effective large-scale milk analysis.
MY	Milk yield.
NEB	Negative energy balance, occurring when energy demand exceeds intake, typically in early lactation.
PCA	Principal component analysis, a method used to account for ancestral groups and population stratification.
PP	Protein percentage.
PPC	Posterior Probability of Causality, assigned to variants during Bayesian fine-mapping.
PS	Population stratification, a potential source of false associations in genetic studies.
PY	Protein yield.
QTL	Quantitative trait loci, specific regions of the genome linked to variation in complex traits.
REML	Restrictive maximum likelihood.
SCC/SCS	Somatic cell score/count, a measure used to target improved udder health.
SIGNET	Simultaneous Inference of Genes and Tissues, a framework that integrates GWAS results with gene regulatory networks.
SNPs	Single nucleotide polymorphisms, the most common type of genetic variation used as markers in GWAS.
SV	Structural variants, which are larger-scale genetic variations than SNPs.

CHAPTER 1
Introduction

1.1 Introduction

Milk production, live weight (LWT), and body condition score (BCS) are all key traits in dairy production systems. Although BCS was not directly evaluated in this study, it provides important biological context for understanding how cows allocate energy between milk synthesis and the maintenance of body reserves (Pryce & Harris, 2006; McCarthy et al., 2007). This balance is central to dairy cow performance. Cows that can produce milk efficiently while maintaining adequate body condition are more likely to remain fertile, healthy, and productive across multiple lactations (Patton et al., 2007; Sun et al., 2025). In contrast, excessive mobilisation of body reserves is associated with poorer reproductive performance, increased health issues, and reduced longevity (Domecq et al., 1997). Understanding genetic factors that influence these traits, both individually and in combination, is therefore essential for improving productivity and sustainability of pasture-based systems (Dechow et al., 2004; Berry et al., 2007; Friggens et al., 2007).

In New Zealand, these relationships are particularly important due to the country's predominantly pasture-based, seasonal dairy production system. Most herds operate under spring-calving, with milk production expected to align closely with seasonal pasture availability. Compared with intensive indoor systems used overseas, New Zealand dairy cows are often exposed to greater variation in feed supply, particularly during early lactation when energy demands are highest. As a result, LWT and BCS play a critical role in monitoring how well cows cope with these challenges. Genetic selection must balance production with traits associated with robustness, the ability to maintain health, fertility, body condition, and performance under challenging conditions, alongside efficiency, and resilience (McClearn et al., 2020; Jagusiak et al., 2023). New Zealand dairy breeding programmes reflect these system pressures using multi-trait selection indices that prioritise overall profitability and sustainability rather than selection on milk volume alone. Breeding objectives need to focus emphasis on milk-solids production, fertility, survival, and functional traits. Live weight and BCS are increasingly recognised as important components of these objectives due to their influence on feed efficiency, reproductive performance, and animal welfare (Handcock et al., 2019). However, selection for increased milk production can still create unfavourable pressures on body reserves, highlighting the need for an increased understanding of the relationships between milk production and functional traits (Sneddon et al., 2015a).

The current national dairy herd is made up of 61.1% Holstein-Friesian x Jersey crossbreds, 23.4% Holstein-Friesians, 7.5% Jersey and 8% is made up of other breeds (DairyNZ and LIC, 2025). During the production season of 2024/2025 there were a total of 4.677 million cows across 10,370 herds in the New Zealand dairy industry, with a production total of 21,012 million litres with 1,937 million kgs of milk-solids processed. In 2024–2025, the nominal average dairy cooperative compensation was $10.75

per kg of milksolids (MS). Milksolids being the combined weight of milk fat and protein produced by the cow or cows. Of which was $1.61 more than the $9.14 five-year average. The national herd test averages for the production season of 2024/2025 and 2023/2024 are presented in Table 1.1.

Table 1. 1: National herd test production averages for New Zealand dairy cattle during the 2023/2024 and 2024/2025 production seasons.

Season	**Milk (litres)**	**Milkfat (kg)**	**Milkfat (%)**	**Protein (kg)**	**Protein (%)**	**Milk-solids (kg)**	**Milk-solids (%)**	**DIM (herd testing)**	**DIM (production)**	**SSC (000 cells/ml)**
2024/2025	4,854	242	5.06	193	4	435	9.06	241	275	157
2023/2024	4,761	234	4.99	189	3.99	423	8.99	241	277	161

Source: DairyNZ and LIC (2025).

Genetic and phenotypic evaluations provide a powerful approach for investigating the genetic architecture of complex traits relevant to dairy production including milk production, LWT, and BCS (Hayes et al., 2010; Hayes & Daetwyler, 2019). These traits are typically polygenic and influenced by many loci of small effect, as well as by environmental and management factors (Hayes et al., 2010; Tam et al., 2019). Importantly, when multiple traits are analysed within the same population, a genetic evaluation can also be used to explore shared genetic influences and potential trade-offs between traits (Jagusiak et al., 2023; Pryce & Harris, 2006). This thesis investigates the genetic basis of daily walk-over-weight LWT and daily milk production traits in New Zealand dairy cattle using pedigree-based animal models and phenotypic analyses. The study utilised detailed phenotypic and pedigree data from a single New Zealand dairy herd, with repeated daily records collected on the same animals across the 2023/2024 lactation season. This allows changes in milk production and live weight to be modelled over time, rather than relying on single time-point measurements that may not fully capture biological dynamics (Li et al., 2018).

CHAPTER 2

Genes associated with Live Weight, Body Condition Scores and Milk traits in Dairy Cattle and the implications they have on Production: A Literature Review

2.1 Introduction

Live weight and BCS are widely used indicators of health, productivity, and welfare in dairy cattle. Live weight reflects size and growth across an animal's lifetime, while BCS provides an estimate of body reserves through a visual muscle-to-fat ratio, which fluctuates with nutrition, season, and physiological state. Traditionally, these traits have been assessed using scales, walk-over weighing systems, or visual and tactile scoring of fat cover over anatomical landmarks. Advances in technology, such as ultrasound measures of fat depth and automated camera-based BCS systems, offer greater precision however their adoption is often limited by cost.

Beyond measurement, LWT and BCS are known to have a significant genetic component, with heritability estimates and genomic studies confirming the role of genes in regulating growth, feed efficiency, fat deposition, and metabolic balance (Alemu et al., 2025; Dechow et al., 2001). Many scientific research papers have found quantitative trait loci (QTL) and candidate genes which are associated with energy balance, weight regulation, and fat mobilisation (Dechow et al., 2001; Krattenmacher et al., 2019; Pryce & Harris, 2006; Yin & König, 2019). These findings highlight the potential of genetic selection for improved MS production, fertility, and resilience. This is particularly relevant in the New Zealand dairy industry, where pasture-based seasonal systems place strong pressure on cows to achieve mating and calving objectives while maintaining metabolic health.

The genes regulating LWT and BCS also have been reported to co-regulate milk traits (Berry et al., 2007; Krattenmacher et al., 2019). This includes but is not limited to volume, persistency and composition (Cohen-Zinder et al., 2005). Milk yield (MY) and composition are fundamental traits for the global dairy industry, which directly influence farmer profitability, processing efficiency, and consumer satisfaction (Miglior et al., 2007; Sneddon et al., 2015b). High-yielding cows influence farm revenue, while the proportions of fat, protein, and lactose determine both product quality and processing value. Processors depend on consistent milk composition for manufacturing efficiency, and consumers increasingly demand products with specific nutritional or sensory profiles. Which is why improving milk yield and composition remains central to dairy breeding and management, linking animal biology to economic and market outcomes (Hayes et al., 2010).

Reliable phenotypic recording underpins progress in these traits. Traditional milk recording programs routinely measure yield, somatic cell count (SCC), and key components across lactations, providing the foundation for genetic evaluation. More recently, mid-infrared spectroscopy (MIR) has expanded the scope of milk analysis. Allowing large-scale, cost-effective prediction of traits such as fatty acid profiles, casein content, and even indicators of energy balance (Ma et al., 2021; Pegolo et al., 2021). Techniques like gas chromatography and mass spectrometry can offer greater biochemical precision, but their

expense limits their use to research settings (Tam et al., 2019). Using a combination of such methods generate rich datasets essential for studying the genetic control of milk traits.

Genetic studies have long established that milk yield and composition are moderately to highly heritable, meaning that genetic improvement through selection is both possible and effective (Hayes et al., 2010; Miglior et al., 2007). However, it is widely established that milk traits are complex, influenced by many genes with small effects and by interactions between genetic and environmental factors (Hayes et al., 2010; Tam et al., 2019). Traditional pedigree-based approaches capture overall additive variance but reveal little about the underlying molecular mechanisms (Bourrat & Lu, 2017; Hayes et al., 2010). The same is also true for both LWT and BCS, which is why utilising genetic selection to improve traits while both mitigating genetic drawbacks is notoriously difficult to manage (Dechow et al., 2001; Pryce & Harris, 2006).

The development of Genome-Wide Association Studies (GWAS) has transformed how these traits are studied (Hayes & Daetwyler, 2019; Ma et al., 2021). By examining hundreds of thousands of single nucleotide polymorphisms (SNPs) across the genome, GWAS identifies specific regions and genes associated with milk yield, fat, and protein composition. These insights have revealed genes and pathways linked to lipid metabolism, mammary gland function, and milk protein synthesis, providing a molecular context for observed phenotypic variation (Cohen-Zinder et al., 2005; Littlejohn et al., 2014). When integrated with transcriptomic and epigenetic data, GWAS helps clarify how gene regulation and expression influence production efficiency and milk quality (Fang et al., 2019; Wu et al., 2017).

This review examines phenotypic, genetic and GWAS approaches in dairy cattle research, with a focus on genes associated with LWT, BCS, and milk traits. The aim is to summarise current understanding of their genetic architecture and biological connections, and to discuss their relevance to dairy production and breeding programmes.

2.2 Importance of Genome-Wide Studies

Genome-Wide Association Studies are central to understanding the complex genetic architecture of traits in dairy cattle (Abdalla et al., 2023; Sousa et al., 2024; Yin & König, 2019). They allow researchers to identify SNPs linked to economically important traits and serve as the foundation for modern genomic selection programs (Abdalla et al., 2023; Alim et al., 2014; Sousa et al., 2024). Before GWAS strategies were developed, researchers used linkage mapping and candidate gene methods to find quantitative trait loci. These older methods had low resolution and were not effective for complex traits like LWT or BCS. The arrival of dense SNP arrays and reference genomes allowed GWAS to study genetic variation across the entire genome with much greater accuracy. By connecting genotype to phenotype, GWAS helps breeders make more precise decisions to improve MY, MS production, health, and longevity

simultaneously (Dechow et al., 2001; Sousa et al., 2024). To achieve reliable results, GWAS must be supported by robust methodology (Sousa et al., 2024; Yin & König, 2019).

2.2.1 Imputation

Imputation is a standard process used in GWAS analysis, referring to the statistical inference of unobserved genotypes from a reference population. Genotype imputation infers missing markers in individuals genotyped on lower-density panels using reference sets from higher-density or whole-genome sequenced animals. This method is a more cost-effective way to increase marker density and GWAS power, enabling large-scale studies of structural variants (SV), and supporting downstream analyses (Chen et al., 2021). Despite its functionality, imputation exhibits reduced accuracy for structural variants and rare alleles. It frequently necessitates the exclusion of potentially informative variants and is highly dependent on the quality and density of reference panels. Simply put, its effectiveness is constrained in regions of high linkage disequilibrium, by methodological processing choices, and by limitations inherent to current sequencing technologies (Chen et al., 2021).

A panel often used in this field is 'The 1000 Bull Genomes Project' which is a whole genome sequence collection from 2703 animals that embodies a significant proportion of international cattle diversity (Hayes & Daetwyler, 2019; Kamprasert et al., 2024). This extensive reference panel enables accurate imputation of SNP array data to whole-genome level, improving the precision of GWAS and the discovery of causal mutations linked to production and fertility traits. By integrating data across breeds, the 1000 Bull Genome Project improves genomic prediction accuracy and supports more informed, data-driven breeding strategies (Hayes & Daetwyler, 2019). The use of imputation from medium-density to high-density SNP panels can enable more accurate identification of quantitative trait loci (Sousa et al., 2024). Medium-density and high-density SNP panels differ mainly in marker density and imputation potential. Medium density arrays (≈50K SNPs) are commonly used for routine genotyping and are imputed to high density or whole-genome sequence data using reference panels like the 1000 Bull Genomes Project to enhance detection of true genetic signals. High density panels (≈630K SNPs) provide greater genomic coverage and greater prediction accuracy, though the added benefit over whole genome sequencing is often small. Implementing hybrid approaches where key whole genome sequence variants are added to medium density panels, is a cost-effective way to improve genomic prediction reliability by up to 4% for production traits (Hayes & Daetwyler, 2019). Crossbreed or multi-population GWAS are useful for finding both shared and breed-specific QTL in different cattle populations. Combining data from multiple breeds increases statistical power and helps pinpoint loci that are conserved or unique. However, differences in linkage disequilibrium, allele frequencies, and trait definitions can make analysis and imputation more difficult. When done carefully, multi-breed GWAS can identify causative variants and support genomic selection across populations with mixed ancestry.

2.2.2 Population Stratification and Cryptic/Family Relatedness

Due to the nature of GWAS being observational studies of genetic traits, unlike experimental models where environment and mating can be controlled, makes GWAS highly susceptible to confounding. Population stratification or PS occur when differences in allele frequencies occur between subpopulations due to non-random mating, often from geographic isolation and genetic drift. Population stratification can produce spurious associations, where marker alleles appear linked to traits simply because both vary across ancestral groups (Hellwege et al., 2017). Thus, corrections such as genomic control (GC) and principal component analysis (PCA) are essential to account for these effects and ensure associations reflect true biological relationships (Price et al., 2010). Correcting for these effects is essential, as failure to do so renders studies unreliable, wasting funding and resources. Especially when the degree of confounding is tightly related to sample size, larger studies are significantly more impacted. Approaches used for correction include the genomic relationship matrix (G-matrix) (Schmidtmann et al., 2023; Yin & König, 2019) or leave-one-chromosome-out (LOCO) models (Yin & König, 2019).

Furthermore, using de-regressed estimated breeding values (dEBVs) (Schmidtmann et al., 2023; Yin & König, 2019) instead of raw phenotypes strengthens statistical precision, especially for traits influenced by management or environmental factors (Yin & König, 2019). While methods such as PCA, GC, and model-based ancestry inference are widely employed, each with their own limitations. This includes incomplete adjustment for cryptic relatedness, residual substructure, sensitivity to marker density, and reliance on accurate population parameters (Hellwege et al., 2017). In addition to PS, cryptic relatedness presents another major confounding factor in GWAS, this refers to hidden genetic relationships between individuals that are not recorded in pedigrees but still influence association statistics (Abdalla et al., 2023; Hellwege et al., 2017; Price et al., 2010). When related animals are treated as unrelated, their genetic similarity can inflate test statistics and create false-positive associations. Standard corrections like PCA and GC don't always capture this fine-scale relatedness, leaving residual bias (Hellwege et al., 2017). Linear Mixed Models (LMMs) that include a G-matrix improve control by modelling how individuals are related across the genome, but their success depends on accurate genomic data and sufficient sample sizes (Mullen et al., 2011; Price et al., 2010).

Even improved methods such as LOCO can overestimate SNP effects in populations with many closely related individuals. In livestock studies, failing to correct for cryptic relatedness has caused severe inflation in test statistics and thousands of false-positive results (Schmidtmann et al., 2023). Yin et al (2019) reported a preliminary GWAS on German Holstein cattle without accounting for any polygenic effects, where more than 2000 significant SNPs were detected even after Bonferroni correction.

However, the genomic inflation factor was 6.04, indicating severe inflation of the test statistics and suggesting that many detected associations were likely false positives arising from unaccounted population structure and relatedness. Thus, properly accounting for cryptic relatedness is crucial to ensure GWAS findings reflect true genetic effects rather than artefacts of shared ancestry. Although GWAS is powerful, it has limitations. Many loci explain only a small part of heritability, and most significant associations are often in non-coding regions, making them hard to interpret (Abdalla et al., 2023; Wu et al., 2017). Genome-wide association study often highlights broad genomic regions rather than specific causal variants, and standard models usually miss rare variants, gene interactions, and gene–environment effects (Bourrat & Lu, 2017; Chen et al., 2021; Hayes & Daetwyler, 2019 ; Wu et al., 2017). Integrating GWAS with functional genomics, transcriptomics, and epigenetic data helps identify biologically meaningful targets.

Bourrat & Lu (2017) discussed the missing heritability problem, which occurs when GWAS estimates of heritability are much lower than those from traditional methods. Traditional approaches often overestimate heritability because they include non-additive genetic effects, shared environmental influences, and assortative mating. Genome-wide association study underestimates heritability because it only captures DNA-based variants and ignores epigenetic effects. The gap is therefore due to both hidden heritable variations, such as epigenetics and rare variants, and the bias in traditional estimates (Bourrat & Lu 2017). If GWAS included epigenetic effects and traditional methods excluded non-additive influences, most of the difference would disappear. Post-GWAS analyses further refine and validate genetic discoveries (Sousa et al., 2024). Fine-mapping techniques, such as Bayesian Fine-Mapping, is a method that identifies independent signals within a QTL region and assigns a Posterior Probability of Causality (PPC) to each variant. It accounts for variant effects, population structure, and phenotypic reliability, and can integrate functional annotations to prioritise biologically important variants. Methods such as these are important for GWAS in dairy genetics because it narrows down associated regions to likely contain causal variants and genes, reducing credible sets, improving precision, and supporting efficient genomic prediction and functional analysis in complex traits (Jiang et al., 2018). Functional annotation through tools like DAVID (Database for Annotation, Visualisation and Integrated Discovery) (Sousa et al., 2024; Maiorano et al., 2018) and enrichment analyses using Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) helps translate these variants into biological meaning (Abdalla et al., 2023; Maiorano et al., 2018). This method can reveal how certain genes influence metabolism, growth, or immune regulation (Abdalla et al., 2023).

To link statistical associations to biological mechanisms, post-GWAS analyses often integrate expression quantitative trait loci (eQTL) data. Mapping of eQTL identifies genetic variants that regulate gene expression, highlighting how SNPs influence phenotypic variation through transcriptional control. Integrating GWAS findings with RNA-Sequence transcriptome data enables validation of candidate

genes (Sousa et al., 2024) through tissue-specific expression providing a clearer link between molecular function and observed phenotypes. Another paper by Wu et al., (2017) supported such integrative approaches, using SIGNET (Simultaneous Inference of Genes and Tissues), to demonstrate that combining GWAS with transcriptome data enables validation of candidate genes utilising tissue-specific expression. As mentioned earlier, while GWAS identifies associated loci with most falling in non-coding regions, SIGNET can address this by integrating GWAS summary statistics with tissue-specific gene regulatory networks. Allowing simultaneous inference of phenotype-associated genes and their relevant tissues within a Bayesian framework (Wu et al., 2017). These networks improve specificity, model dependency (example used by Wu et al., 2017 is the Markov Random Field model), and statistical power for gene prioritisation and tissue identification. SIGNET further validated findings by quantifying tissue-phenotype associations with posterior inclusion probability, successfully highlighting tissues known to be biologically relevant to the trait. Despite these powerful post-GWAS approaches, functional interpretation of associated variants remains challenging (Pereira et al., 2023). Causal variants are difficult to pinpoint due to linkage disequilibrium and the prevalence of non-coding associations, and the nearest gene to a GWAS signal is often not the causal one (Pereira et al., 2023; Tam et al., 2019). Overall, these limitations highlight that post-GWAS functional analyses are invaluable but require careful, context-specific interpretation.

2.2.3 Epistasis, Epigenetic and Environmental factors

Beyond single-gene effects, complex traits in dairy cattle are influenced by interactions between genes, known as epistasis (Alim et al., 2014; Ardicli et al., 2019). These interactions occur when the effect of one gene depends on the presence of other genes, meaning that individual SNP effects may not fully explain phenotypic variation. For example, interactions between leptin (LEP) and growth hormone receptor (GHR), insulin-like growth factor 1 (IGF1) and leptin (LEP), and beta-lactoglobulin (LGB) and growth hormone (GH) have been shown to impact growth and fattening traits, including average daily gain and feed efficiency (Ardicli et al., 2019; Bourrat & Lu, 2017). Ignoring these gene-gene interactions can therefore underestimate the genetic contribution to economically important traits. Epigenetic modifications, such as DNA methylation and chromatin remodelling, also play a role by regulating gene expression without altering the DNA sequence. These heritable modifications contribute to the "missing heritability" in GWAS as standard SNP-based analyses do not capture their integrated effects (Ardicli et al., 2019; Alim et al., 2014; Bourrat & Lu, 2017). Integrating epigenetic data, such as regulatory annotations and tissue-specific marks, has helped identify non-coding variants with functional significance and link molecular mechanisms to phenotypic outcomes. Additionally, environmental factors interact with both genetic and epigenetic mechanisms, influencing trait expression across developmental stages and production systems. Together, epistasis, epigenetics, and environmental

effects highlight the complexity of dairy traits and the need for integrative approaches to fully understand the genetic architecture underlying growth, fat deposition, and metabolic performance.

2.3 Genetic Architecture of Live weight and Body Condition in Dairy Cattle

The genetic architecture of complex traits such as LWT and BCS reflects the number, genomic distribution, and effect sizes of loci influencing biological processes underlying growth and energy balance. The two traits are quintessentially polygenic, with variation controlled by many thousands of loci, most contributing small additive effects. Rather than being governed by single major genes, LWT and BCS emerge from interacting gene networks associated with skeletal development, muscle accretion, lipid metabolism, and endocrine regulation, whose expression is further modulated by production system and physiological state.

2.3.1 Heritability

Heritability estimates vary with age for LWT, measurement method and population. In New Zealand's multi-breed dairy population, the heritability of scale-recorded LWT is consistently high, ranging from 0.57-0.64 indicating strong additive genetic control (Alemu et al., 2025). In contrast, subjective BCS and LWT scores tend to show lower heritability due to evaluator variation and greater environmental sensitivity. Estimates in research herds typically range from 0.19 (Jagusiak et al., 2023) to 0.45 (Dechow et al., 2001), while New Zealand dairy populations show relatively stable heritability throughout different stages of lactation, with a mean estimate of approximately 0.26 (Pryce & Harris, 2006). Producer-recorded data often produce lower estimates (0.07–0.20), due to marginal error. These values indicate that while management plays a major role, sufficient genetic variation exists to enable selection for improved body condition. This contrast emphasises this important biological distinction: live weight largely reflects cumulative growth and skeletal development and BCS reflects dynamic, reversible energy storage responding rapidly to nutritional and metabolic pressures.

2.3.2 Quantitative Trait Loci

Genome-wide association studies and QTL mapping have demonstrated that genomic regions influencing LWT are distributed across the bovine genome, with notable effects on chromosomes such as BTA 14, BTA 5, and BTA 18 (Chen et al., 2021; Schmidtmann et al., 2023; Yin & König, 2019). The major locus on BTA 14, frequently associated with stature, carcass weight, and milk production, illustrates the pleiotropic nature of growth regulation. Similarly, regions on BTA 5 and BTA 18 contribute to both structural size and functional traits such as calving difficulty, reinforcing that selection on LWT cannot be isolated from correlated biological consequences. These findings indicate that genetic control of body size is tightly integrated with reproductive physiology and the physiological mechanisms underlying milk production, rather than operating independently. In contrast to LWT, BCS represents a

more metabolically flexible trait reflecting short-term energy balance. Moderate heritability estimates indicate that while genetic differences exist, environmental factors such as nutrition, health, and lactation stage dominate phenotypic expression. QTL for BCS cluster in genomic regions related to lipid metabolism and endocrine regulation, notably on BTA 5, BTA 6, and BTA 14 (Ardicli et al., 2019; Schmidtmann et al., 2023). The co-location of BCS QTL with udder depth, dairy character, and fat-related traits illustrates the interconnected nature of body reserves, conformation, and production.

2.3.3 Candidate genes

At the gene level, candidate loci such as LCORL, NCAPG, IGF1, and fibroblast growth factors (e.g. FGF6, FGF23) highlight the importance of the somatotropic axis and skeletal-muscle development pathways (Jiang et al., 2018; Maiorano et al., 2018; Yin & König, 2019). The repeated identification of LCORL and NCAPG across breeds and species indicates evolutionary conservation of growth regulation mechanisms. However, despite the identification of these biologically plausible genes, most of the genetic variance in LWT remains explained by numerous small-effect loci, supporting the use of genomic selection approaches rather than marker-assisted selection targeting individual QTL. Candidate genes such as ABCC9, NPFFR2, TMEM117, and appetite regulators like MC4R emphasise the central role of energy homeostasis pathways in determining BCS (Jiang et al., 2018; Schmidtmann et al., 2023; Sousa et al., 2024). Unlike LWT genes that largely affect structural size, BCS genes influence fat deposition, mobilisation, and metabolic signalling (Schmidtmann et al., 2023; Zhang et al., 2023). Importantly, research indicates that the lower boundary of BCS may be under stronger genetic control than greater condition levels, implying that genetic differences become most apparent under metabolic stress, such as early lactation negative energy balance (Friggens et al., 2007).

2.3.4 Polygenic nature and gene networks

An emerging theme in the literature is that the genetic architecture of LWT is developmentally dynamic, different sets of genes contributing to growth at different life stages (Chen et al., 2021; Yin & König, 2019). Maternal genetic effects appear particularly important for birth weight and early growth, whereas direct genetic effects become increasingly important during later skeletal and muscular development (Yin & König, 2019). This temporal regulation implies that selection strategies may affect juvenile growth, mature size, and metabolic load differently, and therefore should consider age-specific genetic parameters rather than treating LWT as a single static trait (Chen et al., 2021; Sousa et al., 2024 ; Yin & König, 2019). The polygenic nature of BCS also reinforces the need for genomic prediction rather than reliance on major genes (Chen et al., 2021). Pathway analyses indicate involvement of oxidative phosphorylation, fatty acid degradation, and insulin signalling networks, indicating that BCS is fundamentally a reflection of whole-body metabolic efficiency rather than simple fat storage (Zhang et al., 2023). The involvement of oxidative phosphorylation, fatty acid degradation, and insulin signalling

pathways supports the idea that selection for BCS may indirectly improve robustness, health, and fertility by altering how cows allocate and utilise energy during periods of high milk production and negative energy balance.

2.3.5 Influence of environment on genetic expression

While LWT is strongly heritable, its phenotypic expression is constrained by genotypic and environmental interactions. Production system differences, such as pasture-based versus total mixed ration systems, alter the realised expression of genetic potential for body mass (Fulkerson et al., 2008). Furthermore, homeorhetic control during early lactation prioritises nutrient partitioning towards milk production, leading to genetically regulated mobilisation of body reserves that cannot be fully compensated by increased feed intake (Bauman & Currie, 1980; Friggens et al., 2007). This supports the concept that genetic regulation of energy allocation is not merely responsive to environment but is actively programmed, particularly in high-producing cows. In New Zealand's pastoral system, the genetic correlation can change from positive in early lactation (+0.16) to negative in late lactation (−0.25) (Horan et al., 2005; Pryce & Harris, 2006). This indicates that cows with greater BCS in early lactation have the reserves necessary to sustain production later in the season when pasture availability may be limited.

2.4 Genetic Architecture of Milk Traits in Dairy Cattle

Similarly, for both LWT and BCS, the genetic architecture of milk production traits is multifaceted in dairy cattle and is characterised by a polygenic background (Hayes et al., 2010; Hayes & Daetwyler, 2019). Genetic analyses have been invaluable in synthesising knowledge regarding these genetic factors, allowing the identification of key genes and genomic regions that contribute to variation in milk yield, composition, and udder health. A recurrent finding is that two key genes consistently emerge as major contributors across multiple milk production traits; DGAT1 located on chromosome 14 and ABCG2 on chromosome 6, which influence milk volume, fat and protein percentages (Fang et al., 2019; Si et al., 2024). In addition to DGAT1 and ABCG2, other genes discussed in sections 2.4.1 and 2.4.2 later in this review, including GHR, MGST1, PAEP and the casein gene cluster, further illustrate the pleiotropic effects of genes on both milk production and milk composition traits in dairy cattle.

2.4.1 Milk Yield

Milk yield, represents the primary driver of economic output in dairy systems and along with LWT and BCS, is under complex polygenic control (Pryce &Harris, 2006). The genetic architecture of milk yield is characterised by a highly polygenic and leptokurtotic structure, where thousands of loci contribute small additive effects alongside a limited number of major QTL with disproportionately large influence on phenotypic variance (Blott et al., 2003; Hayes et al., 2010; Hayes & Daetwyler, 2019). Heritability

estimates for milk yield are generally moderate, indicating substantial scope for genetic improvement. Across dairy populations, SNP-based heritability for milk yield and energy-corrected milk (ECM) typically ranges from 0.14 to 0.39 (Hayes et al., 2010; Schmidtmann et al., 2023). However, greater estimates have been reported in controlled research herds where environmental noise is reduced (often exceeding 0.50 for ECM), indicating that more than half of the observed variation in ECM can be attributed to genetic differences among animals (Chen et al., 2021; Krattenmacher et al., 2019). However, the genetic control of milk yield is not static across lactation. Genetic correlations between early-, mid-, and late-lactation yield remain positive but decline with increasing DIM (Days In Milk), indicating that partially distinct biological mechanisms regulate yield as cows transition from mobilisation of body reserves to sustained nutrient intake (Friggens et al., 2007; Haile-Mariam & Pryce, 2017 ; Krattenmacher et al., 2019).

This dynamic relationship creates an important link between MY, LWT, and BCS, as selection for high early-lactation yield is genetically antagonistic with energy balance and body condition recovery (Patton et al., 2007; Pryce & Harris, 2006). Genome wide association and fine-mapping studies consistently identify several major loci underlying variation in milk yield (Hayes et al., 2010). DGAT1 on BTA14 remains the most prominent QTL, where the K232A mutation alters lipid synthesis and partitioning, producing pleiotropic effects on yield and composition (Grisart et al., 2002). While certain alleles increase fat synthesis, they often reduce total milk volume, illustrating the trade-off between secretion concentration and liquid output. Similarly, ABCG2 on BTA6 harbours a causative Y581S mutation that significantly reduces milk yield while increasing fat and protein percentages, reinforcing the biological antagonism between volume and component density (Cohen-Zinder et al., 2005). The growth hormone receptor gene (GHR) on BTA20 also exerts strong effects on yield through endocrine regulation of nutrient partitioning toward the mammary gland (Blott et al., 2003; Schmidtmann et al., 2023). A defining feature of milk yield architecture is its near-perfect genetic linkage with lactose synthesis. Lactose acts as the primary osmotic agent of milk secretion, drawing water into mammary secretory vesicles and thereby determining milk volume (Neville, 2001). Consequently, genetic correlations between milk yield and lactose yield are typically very high (Lopdell et al., 2017; Haile-Mariam & Pryce, 2017). Genes involved in lactose biosynthesis and transport, such as LALBA (encoding α-lactalbumin, a key subunit of lactose synthase), directly regulate liquid milk output rather than simply component mass (Hayes & Daetwyler, 2019). Because lactose acts as the primary osmotic regulator of milk secretion and therefore influences milk volume, loci affecting lactose synthesis can also affect the concentration of other milk constituents (Neville, 2001). This mechanism helps explain why many loci affecting MY also influence dilution of fat and protein percentages, forming the biological basis for the yield–composition trade-offs discussed in Section 2.4.2.

2.4.2 Milk Composition

Fat Percentage/Yield

While milk yield drives economic productivity, milk composition drives product value and processing efficiency (Alim et al., 2014; Auldist et al., 2007). Like MY, milk composition traits are generally highly heritable. Differences in fat and protein percentages between breeds contributes to differences in product yield and processing values across dairy markets (Hayes et al., 2010). Jersey-type breeds, for example, are well known for producing much creamier (greater fat percentage) milk than Holstein-Friesians, which have greater milk volume characteristics (Lembeye et al., 2016). In recent years, specialty or niche milk products have become popular and the demand for milk which are low fat, high in protein, low in lactose are some examples of growing niche markets and attract increased value (Alim et al., 2014; Sneddon et al., 2016). Given the importance of the dairy industry to New Zealand's economy, variation in consumer and processing demand must be considered when evaluating milk composition traits. According to the 2023/2024 NZ dairy cow statistics, ~60% of New Zealand's herd are Holstein-Friesian x Jersey Crossbred, with a further 23.9% of the herd being Holstein-Friesian and 7.5% Jersey (DairyNZ, 2025). The predominance of crossbred animals reflects the long-term adoption of crossbreeding within New Zealand pasture-based dairy systems (Lembeye et al., 2016). Si (2024) identified 36 QTLs associated with fat percentage (FP) in a meta-analysis. Fortunately, FP is highly heritable and has been found to be as high as 0.83 in one study (Hayes et al., 2010). A few key genes, DGAT1, ABCG2 and AGPAT6, have been identified as major contributors and are thought to have some of the largest impacts on milk fat production and composition (Cohen-Zinder et al., 2005; Grisart et al., 2002; Jiang et al., 2018; Littlejohn et al., 2014).

Protein Percentage/Yield

The genetic architecture of protein percentage (PP) in dairy cattle is characterized by a polygenic structure in which a small number of major genes exert relatively large effects alongside a large number of loci with small additive contributions (Hayes & Daetwyler, 2019). Unlike protein yield (PY), which is strongly influenced by milk volume, PP reflects both protein synthesis capacity and dilution processes and therefore often exhibits higher heritability estimates and stronger signals from major-effect loci (Haile-Mariam & Pryce, 2017; Hayes et al., 2010). Heritability estimates for PP are typically high, often ranging from 0.45-0.80 depending on breed, population, and statistical model (Lopdell et al., 2017; Si et al., 2021). In contrast, PY heritability is low to moderate ranging from 0.14-0.25 (Pegolo et al., 2021; Si et al., 2021). This substantial additive genetic variance makes PP a stable and effective target for genetic selection aimed at improving milk quality and processing characteristics. Genome-wide association studies consistently identify several key genomic regions acting as primary regulators of protein concentration. One of the most influential loci is DGAT1 on BTA14, where a missense mutation

(K232A) has major effects on milk composition. The 232A allele is associated with a reduction in PP (approximately −0.26%) as reported by Si et al (2021), while simultaneously increasing milk volume, illustrating a classic pleiotropic trade-off between yield and concentration.

Another major locus is ABCG2 on BTA6, where a tyrosine-to-serine substitution (Y581S) has been identified as the causative mutation underlying a prominent QTL affecting milk production traits (Cohen-Zinder et al., 2005). Cohen-Zinder et al. (2005) observed the Y581 allele increases protein and fat concentration (approximately 0.13 and 0.16 percentage points, respectively) but decreases overall milk yield, again demonstrating antagonism between component concentration and volume. The gene MGST1 on BTA5 has also been repeatedly implicated across breeds and populations as a significant contributor to both milk yield and PP, demonstrating involvement in broader metabolic regulation of mammary secretion (Si et al., 2024). On BTA11, PAEP, which encodes β-lactoglobulin, directly affects whey protein composition and therefore contributes materially to overall PP and processing quality (Ardicli et al., 2019; Hayes & Daetwyler, 2019). Similarly, variants in AGPAT6 on BTA27 show pleiotropic effects, where increased expression is associated with greater PP but reduced PY and milk volume (Littlejohn et al., 2014). This further supports the concept that PP is controlled by regulatory trade-offs rather than simple increases in protein synthesis alone. The genes encoding the four major caseins (κ-casein, β-casein, αS1-casein, and αS2-casein) are physically clustered within a ~250-kb region on BTA6, forming one of the most important genomic regions affecting milk protein composition (Alim et al., 2014; Pegolo et al., 2021). This tight clustering facilitates coordinated regulation but also introduces competitive dynamics among fractions. Genomic analyses reveal strong negative correlations between certain protein fractions, for example between β-casein and αS1-casein (approximately −0.71), which may reflect competition for transcriptional machinery, amino acid supply, or secretory capacity within mammary epithelial cells (Pegolo et al., 2021). Bayesian network and causal modelling approaches further indicate that κ-casein and β-casein act as regulatory "root nodes," directly or indirectly influencing the concentrations of other protein fractions (Jiang et al., 2018; Pegolo et al., 2021). This highlights that PP is structured through hierarchical regulatory networks rather than independent synthesis of individual proteins.

2.5 Cross-Trait Relationships

The genetic architecture of dairy cattle is defined not only by the individual inheritance of LWT, BCS, and milk traits, but by the intricate network of biological relationships that connect them. These cross-trait relationships arise from shared physiological pathways governing energy partitioning, endocrine regulation, mammary gland function, and structural development. Understanding these relationships is central to interpreting GWAS findings and to designing breeding objectives that avoid unintended antagonistic responses. A consistent finding across international populations is the moderate negative genetic correlation between MY and BCS (approximately −0.28 to −0.34) (Dechow et al., 2001; Pryce

& Harris, 2006; Schmidtmann et al., 2023). Thus reflecting a biological prioritisation of nutrient allocation toward lactation at the expense of body tissue reserves. Cows with high genetic merit for yield tend to partition energy toward the mammary gland, particularly in early lactation when intake capacity lags milk energy output (Bauman & Currie, 1980; Friggens et al., 2007). This antagonism is most clear during negative energy balance (NEB). Genetic correlations between milk energy output and energy balance range from −0.26 to −0.59 (Krattenmacher et al., 2019), indicating that selection for greater production intensifies early-lactation metabolic stress. In contrast, dry matter intake (DMI) shows strong positive genetic correlations with energy balance (0.71 to 0.81), displaying intake capacity partially buffers production-driven deficits (Krattenmacher et al., 2019). In pasture-based systems such as New Zealand, the genetic correlation between BCS and fat/protein yield has been reported as positive in early lactation (+0.16) but negative by late lactation (−0.25) (Pryce & Harris, 2006). This temporal shift implies that cows with greater early reserves can sustain production under pasture limitation, but prolonged mobilisation reduces condition later in lactation. Thus, BCS is not merely representative of fat storage, but a dynamic indicator of genetically regulated energy resilience.

Milk yield is almost perfectly genetically correlated with lactose yield (≈0.99) (Lopdell et al., 2017). This relationship reflects the role of lactose as the primary osmotic regulator of milk secretion, which largely determines milk volume (Neville, 2001). Lactose synthesis establishes the osmotic gradient within mammary epithelial cells, drawing water into secretory vesicles and thereby controlling milk volume (Neville, 2001). Consequently, genes influencing lactose metabolism directly affect total milk yield (Hayes & Daetwyler, 2019; Lopdell et al., 2017). However, genetic correlations between milk yield and fat or protein percentages are typically near zero or negative, illustrating the well-established "dilution effect" (Haile-Mariam & Pryce, 2017; Hayes et al., 2010). As milk volume increases, component concentrations often decreases unless synthesis of solids increases proportionally (Haile-Mariam & Pryce, 2017; Neville, 2001). This relationship creates a compositional trade-off central to dairy breeding. Several major pleiotropic loci illustrate this antagonism;

DGAT1 (BTA14): The K232A mutation increases FP and yield but reduces total milk volume and protein yield (Cohen-Zinder et al., 2005; Littlejohn et al., 2014; Zhou et al., 2019).

ABCG2 (BTA6): The Y581 allele increases protein and fat concentrations by approximately 0.13 and 0.16, respectively, while decreasing overall milk yield by 341kg (Cohen-Zinder et al., 2005).

AGPAT6 (BTA27): Variants within AGPAT6 enhances fat synthesis and protein percentages but are associated with reduced milk volume and lactose yield (Littlejohn et al., 2014).

These loci demonstrate antagonistic pleiotropy, where enhancement of component density is achieved at the expense of fluid output. From a systems perspective, this reflects the biological constraint that mammary secretory capacity, substrate supply, and osmotic regulation are interdependently rather than

independently adjustable. Live weight is positively associated with milk yield phenotypically, as larger cows generally possess greater mammary capacity and feed intake potential. However, genetically increased LWT is tightly linked to greater maintenance energy requirements (Alemu et al., 2025). This creates a trade-off between milk production capacity and feed efficiency. Major growth-related loci such as GHR, along with LCORL and NCAPG, influence both stature and milk production, highlighting pleiotropy between structural size and lactation performance (Hayes et al., 2010; Schmidtmann et al., 2023). Selection for larger body size may therefore indirectly increase milk yield while simultaneously elevating metabolic burden. This trade-off is particularly relevant in pasture-based systems, where intake is constrained by forage availability rather than ration formulation. Larger cows require proportionally more energy for basal metabolism, reducing net feed conversion efficiency into milksolids (Horan et al., 2005). Thus, the genetic linkage between LWT and MY must be interpreted within the environmental context in which expression occurs. Fertility traits are strongly influenced by energy balance and body reserves.

Genetic correlations between BCS and fertility traits such as calving rate within 42 days (CR42) range from 0.43 to 0.50 (Patton et al., 2007; Pryce & Harris, 2006), confirming that cows genetically predisposed to maintain body reserves also exhibit improved reproductive performance. Excessive mobilisation of LWT and BCS postpartum delays resumption of ovarian cyclicity and reduces conception rates (Alawneh et al., 2012). Elevated IGF-1 concentrations in early lactation are positively associated with conception success, linking endocrine regulation to both metabolic status and reproductive outcome (Mullen et al., 2011; Patton et al., 2007). Historically, intense selection for milk yield without concurrent emphasis on robustness contributed to declining fertility in many Holstein populations (Dechow et al., 2001). Additionally, greater production and lower BCS are genetically associated with increased susceptibility to metabolic disorders such as ketosis and displaced abomasum with genetic correlations ranging between 0.11–0.22 (Schmidtmann et al., 2023).

These findings reinforce that milk yield cannot be biologically separated from systemic metabolic regulation. Genome-wide analyses reveal overlapping QTL for LWT, BCS, milk traits, and fertility on chromosomes such as BTA5, BTA6, and BTA14 (Ardicli et al., 2019; Schmidtmann et al., 2023). For example: ABCC9 is associated with fat and PY, udder depth, and daughter pregnancy rate. Genes such as NPFFR2 influence body conformation and metabolic stability. These shared loci provide molecular evidence that production, conformation, and metabolic resilience are governed by interconnected regulatory pathways rather than independent genetic modules. Epistatic interactions further complicate interpretation, as the phenotypic expression of one gene depends on the genetic background in which it occurs. Epigenetic modifications and gene–environment interactions also contribute to observed correlations, potentially explaining aspects of the “missing heritability” described by Bourrat & Lu (2017). Thus, cross-trait relationships arise not only from additive pleiotropy but also from regulatory

networks and environmental modulation. From a quantitative genetics perspective, these relationships illustrate the principle that selection on one trait inevitably shifts the equilibrium of others. Therefore, sustainable dairy improvement requires multi-trait genomic evaluation frameworks that recognise biological trade-offs rather than pursuing single-trait maximisation.

2.6 Current Systems

New Zealand's dairy industry operates under a predominantly pasture-based, seasonal calving system that is distinct from the total mixed ration and housed systems common in North America and Europe. Approximately 95% of the national herd is managed on grazed pasture, with minimal reliance on imported feed relative to other major dairy-exporting nations (DairyNZ & LIC, 2025). This production model shapes both the biological demands placed on dairy cattle and the direction of national breeding objectives. Seasonal spring calving is a defining feature of the system. In New Zealand pasture-based systems, cows are typically managed to calve within a compact 6 to 12 week period to align peak lactation with maximum pasture availability and growth (DairyNZ, 2025). Consequently, animals must rapidly transition into early lactation, reach peak milksolids production within a short timeframe, and reconceive within approximately 83 days to maintain a 365-day calving interval (Alawneh et al., 2012). This creates intense physiological pressure during early lactation, particularly in terms of energy balance and body reserve mobilisation (Friggens et al., 2007; Patton et al., 2007). Due to feed supply being pasture-driven and how it fluctuates seasonally, New Zealand cows must exhibit strong metabolic adaptability (Pryce and Harris, 2006). Unlike confinement systems where dietary energy can be increased to support extremely high milk volumes, pasture-based cows must balance intake limitations with production demands. Therefore, absolute milk volume is less economically important than milksolids efficiency, where fat and protein yield is relative to body size and feed intake (Auldist et al., 2007; Lembeye et al., 2016).

This reinforces the importance of cross-trait relationships between LWT, BCS, and milk traits discussed previously. Herd composition reflects this system-driven selection pressure. According to recent LIC 2025 statistics, approximately 60% of the national herd consists of Holstein-Friesian x Jersey crossbreds, with ~24% pure Holstein-Friesian and ~7–8% pure Jersey. Crossbreeding has been widely adopted to exploit heterosis, particularly for fertility, survival, and robustness under pasture constraints (Sneddon et al., 2016). Jerseys contribute elevated milksolids concentration and lower mature LWT, while Holstein-Friesians contribute greater milk volume and stature. The crossbred animal represents a biological compromise optimised for efficiency rather than maximal output. Another defining feature of the New Zealand system is the economic penalty associated with excessive LWT. Larger cows incur greater maintenance costs without proportional increases in milksolids under pasture-limited intake (Hardie et al., 2017; Sieber et al., 1988). As a result, there has been long-standing selection pressure toward moderate body size combined with high component yield and strong reproductive performance.

2.7 Breeding Objectives

New Zealand's national dairy breeding objective is explicitly centred on maximising the efficiency with which cows convert feed into farmer profit under pasture-based, seasonal-calving conditions (DairyNZ, 2025). This objective is implemented through the BW (Breeding Worth) index, a bioeconomic selection index that expresses genetic merit in dollar terms and ranks animals across breeds, herds, and ages according to their expected ability to produce profitable replacement daughters. Unlike breeding programs that emphasise absolute production, the New Zealand system prioritises net efficiency, incorporating both revenue-generating and cost-associated traits into a single optimisation framework (DairyNZ & LIC, 2025; Spelman & Garrick, 1997).

A major structural shift in recent decades has been the increased use of Holstein-Friesian × Jersey crossbreeding, with crossbreds now comprising approximately 60.4% of the national herd. This transition reflects strategic exploitation of heterosis to improve fertility, survival, and metabolic robustness while maintaining strong milksolids production (McClearn et al., 2020; DairyNZ & LIC, 2025; Vance et al., 2013). The BW index comprises ten economically weighted traits categorised into milk production efficiency and robustness components. Economic weights are derived from national bioeconomic modelling and reflect marginal contribution of each trait to farm profit under typical New Zealand dairy industry conditions, including prevailing milk prices, production costs, and pasture-based seasonal-calving systems (DairyNZ & LIC, 2025).

From 16 May 2025, DairyNZ's updated statistics on economic values further reinforce the emphasis on milksolids rather than fluid volume. Milk protein now receives the largest positive economic weight (+$7.10/kg), followed by milkfat (+$5.27/kg). This reflects New Zealand's payment system based on kilograms of milksolids, aligning genetic selection directly with export-driven processing markets where solids underpin revenue (DairyNZ & LIC, 2025; Spelman & Garrick, 1997). Milk volume continues to carry a negative economic weight (−$0.115/kg), representing a deliberate deviation from volume-focused breeding systems (Sneddon et al., 2015b). Excessive milk volume increases processing and transport costs without proportional financial return, particularly in milk powder manufacturing. The negative weighting discourages dilution of milksolids and promotes biological efficiency in component production (Miglior et al., 2007; Sneddon et al., 2015b).

Live weight also remains negatively weighted (−$1.68/kg), acknowledging its strong association with maintenance energy requirements. Larger cows require greater feed inputs to sustain basal metabolism, reducing net conversion efficiency of pasture into milksolids (Berry & McCarthy, 2021; Sieber et al., 1988). Given the positive genetic correlation between body size and feed intake, penalising LWT within BW counterbalances correlated responses that would otherwise increase system costs, reflecting the trade-off between production capacity and maintenance burden (Berry & McCarthy, 2021;

Schmidtmann et al., 2023). Robustness traits capture functional longevity and biological resilience, which are essential in seasonal systems. Fertility now carries a positive weight of +$5.90 per 42-day calving rate (CR42) unit. This metric measures the proportion of cows that calve within the first 42 days of the calving season, recognising that reproductive performance underpins compact calving patterns and sustained lifetime productivity. Functional survival (+$1.90/%) reduces involuntary culling and spreads replacement costs over more lactations (DairyNZ, 2025). Somatic cell score retains a strong negative economic weight (−$46.43/score), targeting improved udder health and reduced mastitis incidence. This substantial penalty reflects both direct treatment costs and indirect losses through reduced milk quality and processor penalties (Berry et al., 2007).

Body condition score receives one of the highest positive economic weights (+$174.60/score), underscoring its central role in energy balance and metabolic resilience. Cows that maintain adequate condition through lactation are better able to sustain production, re-conceive within the restricted mating window, and avoid metabolic disorders (Dechow et al., 2001; Pryce & Harris, 2006; Sun et al., 2025). This weighting directly addresses antagonistic genetic correlations between milk yield and body reserves, embedding energy partitioning biology within the selection objective (Buckley et al., 2000; Krattenmacher et al., 2019).

Additional traits such as gestation length (−$1.95/day) and udder overall (+$1.90/unit) further refine adaptation to seasonal calving and reduce structural culling risk (DairyNZ, 2025). The implementation of a balanced, multi-trait breeding objective has resulted in measurable productivity gains despite reductions in national cow numbers. Milksolids production per cow remains near record levels, while average days in milk remain high relative to historical norms. These improvements reflect cumulative genetic progress in efficiency, fertility, and robustness rather than expansion in herd size (DairyNZ & LIC, 2025). From a quantitative genetics' perspective, the New Zealand breeding objective represents an optimisation strategy under biological and economic constraints. By integrating production traits with LWT, body condition, fertility, health, and structural soundness into a single economic index, the BW framework internalises the trade-offs inherent in energy partitioning (Berry & McCarthy, 2021; Friggens et al., 2007; Pryce & Harris, 2006). The resulting national herd is not selected for maximal output per se, but for sustainable efficiency within a pasture-limited, seasonal system (Spelman & Garrick, 1997; Vance et al., 2013).

2.8 Implications for Breeding Programs

The integration of genomic technologies into dairy breeding has transformed genetic improvement from pedigree-based prediction to molecularly informed selection. Genome-wide associations studies,

genomic prediction, and large reference populations such as the 1000 Bull Genomes Project have substantially increased the resolution at which production, fertility, and body reserve traits can be analysed, improving the detection of QTL and the accuracy of genomic evaluations (Hayes & Daetwyler, 2019; Sousa et al., 2024). However, increased marker density also heightens susceptibility to population stratification and relatedness, highlighting the importance of robust analytical approaches such as linear mixed models and leave-one-chromosome-out (LOCO) methods (Sousa et al., 2024, Yin & König, 2019).

Most dairy traits exhibit a highly polygenic architecture, with additive genetic variation distributed across many loci of small effect (Hayes et al., 2010; Hayes & Daetwyler, 2019). Although major genes such as DGAT1 and ABCG2 have important effects on milk composition, long-term genetic gain depends on capturing cumulative genome-wide variation rather than selecting for individual loci alone (Cohen-Zinder et al., 2005; Hayes & Daetwyler, 2019). Furthermore, pleiotropic effects create biological trade-offs among economically important traits. Alleles associated with increased milk solids concentration often reduce milk volume, while high genetic merit for milk production is frequently associated with reduced body condition and impaired energy balance (Cohen-Zinder et al., 2005; Haile-Mariam & Pryce, 2017).

These relationships demonstrate that sustainable genetic improvement requires multi-trait selection frameworks that account for biological and economic trade-offs. In New Zealand, this principle is reflected in the BW index, which places positive economic weight on milk solids, fertility, and body condition while penalising excessive live weight and milk volume relative to system requirements (DairyNZ & LIC, 2025). By integrating genomic information with bioeconomic objectives, modern breeding programmes increasingly seek to balance milksolids production, feed efficiency, fertility, health, and resilience rather than maximise individual traits in isolation.

CHAPTER 3

Methods and Data

3.1 Data

3.1.1 Farm System and Scale

Massey University Dairy Farm 4 encompasses a total land area of 250 hectares, with an effective grazing platform of 221 hectares. The farm operates under a pasture-based, seasonal calving system typical of New Zealand dairy production (DairyNZ & LIC, 2025). The predominant pasture species are perennial ryegrass (*Lolium perenne* L.) and white clover (*Trifolium repens* L.), with strategic use of supplementary feed (maize silage (*Zea mays* L.), pasture silage) to maintain feed supply and support lactation performance. During Summer chicory (*Cichorium intybus* L.) is utilised as a forage crop, while maize is grown for silage (5–10 hectares grown on-farm), grass silage, baleage, bypass feeds, and in-shed concentrates are incorporated as required. Infrastructure supporting the production system includes a 600-cow capacity feed pad and a 200-cow free stall barn equipped with individual stalls and designated calving areas. Approximately 90% of the farm is artificially drained using tile and mole drainage systems due to the Tokomaru and Ohakea Silt Loam soils, which exhibit poor natural drainage and seasonal moisture variability. The grazing platform operates under a 10-year pasture renewal rotation. Selected stock classes, including weaned heifer calves, rising one-year heifers, and approximately 150 mixed-age dry cows, are grazed off-farm to optimize feed allocation for the milking herd.

3.1.2 Herd Characteristics

The milking herd consists of Holstein-Friesian × Jersey crossbred cows with 100% recorded ancestry, enabling accurate pedigree tracing and genetic evaluation. These cows are milked on a twice-a-day (TAD) system. At the time of data collection, the herd had a BW of 102 (reliability 61%) and a PW of 105 (reliability 71%), indicating above-average genetic merit under national evaluation standards. The farm maintains a stocking rate of approximately 2.7 cows per hectare. During the 2023/2024 production season used in this analysis, peak cow numbers reached 591 with 401kgMS/cow, with total production of 237,205 kg of milksolids and over 2.6 million litres of milk. Mating is conducted using 100% artificial breeding, with no natural mating used. The mating period typically commences in mid-October and continues for approximately 9–10 weeks, consistent with a compact seasonal calving objective.

3.1.3 Data Recording and Monitoring Systems

Massey University Dairy Farm 4 incorporates automated monitoring technologies to ensure accurate and high-frequency data capture. Cows are fitted with Nedap activity collars, which provide behavioural and physiological monitoring data. Milk yield data are collected through LIC Automation YieldSense technology integrated within the milking system, allowing individual cow-level production recording. Milking is conducted across two platforms: a 60-bale Waikato Milking Systems Centrus rotary platform

used for commercial milking and a separate 28-bale Centrus rotary platform dedicated to research activities. The research platform enables controlled data collection while maintaining commercial production conditions.

3.1.4 Data Format

Phenotypic data were structured in a longitudinal format, integrating repeated milk production records with static animal demographic information. Raw datasets were supplied in separate Microsoft Excel (.xlsx) files corresponding to individual recording days and collection periods. These files were consolidated into a single master spreadsheet in which daily production records could be viewed and filtered by animal identification, date, and stage of lactation. Animal-specific information, including age at calving and body weight, was stored on a neighbouring worksheet within the same file and linked to production records using Animal Birth ID as a unique relational identifier. This approach ensured consistent alignment between production traits and corresponding animal-level covariates. An example of data structure is presented in Table 3.1

Table 3. 1: Representative sample of cleaned production records used in the analysis.

Cow ID	Collection Date	DIM	Milk Yield (kg/day)	Fat (%)	Protein (%)	Live weight (kg)	Milksolids (kg/day)	Age at Calving (years)	Breed
CBYC-11-273	29AUG2023	6	27.5	5.8	4.2	581.5	2.71	12	HF
CBYC-14-21	29AUG2023	12	26.3	4.2	4.0	486.8	2.16	9	HFxJ
CBYC-14-57	16OCT2023	63	11.9	6.3	4.3	595.8	1.26	9	HFxJ
CBYC-14-7	12SEP2023	0	17.3	6.5	4.9	483.3	2.44	9	HFxJ
CBYC-15-37	15AUG2023	9	24.9	6.5	3.8	585.6	2.75	8	HF
CBYC-15-6	08AUG2023	22	21.9	5.9	4.3	586.7	2.25	7	HFxJ
CBYC-16-13	05SEP2023	3	19.8	6.2	4.9	526.5	2.20	7	HF
CBYC-16-47	22AUG2023	7	18.6	5.9	4.8	605.1	2.02	7	HFxJ
CBYC-16-52	08AUG2023	14	18.1	5.6	4.1	565.9	1.62	6	HFxJ
CBYC-18-18	08AUG2023	20	10.2	10.5	4.9	473.2	1.57	4	J
CBYC-18-45	08AUG2023	12	31.4	4.4	4.1	474.3	2.57	4	HFxJ
CBYC-18-5	08AUG2023	6	28.3	6.6	4.5	548.7	3.16	5	HF
CBYC-19-30	08AUG2023	13	29.0	5.3	4.4	587.5	2.88	3	HFxJ
CBYC-19-64	08AUG2023	14	24.1	4.1	4.3	565.8	2.04	3	HF
CBYC-20-26	12SEP2023	9	23.1	5.0	4.1	546.2	2.23	3	HF
CBYC-20-56	29AUG2023	11	22.3	4.0	3.7	507.7	1.77	3	HF
CBYC-20-59	08AUG2023	15	22.2	4.5	4.2	571.3	1.94	2	HF
CBYC-20-7	08AUG2023	13	13.3	6.0	4.1	565.4	1.37	3	HFxJ
CBYC-21-50	24OCT2023	93	20.6	4.6	4.0	363.8	1.82	1	HF
CCRK-14-101	29AUG2023	8	24.9	4.5	3.9	517.8	2.13	9	HF

The final analytical dataset comprised both categorical variables and continuous variables describing physiological performance, including milk yield (kg/day), milk composition traits (fat and protein kg and %), LWT (kg), and DIM (days). This longitudinal structure was essential for the implementation of linear mixed models, as it allowed appropriate handling of repeated measurements obtained from the same animals over time.

3.2 Methods

3.2.1 Compiling Data

Milk production and LWT data were provided as separated daily excel files corresponding to individual recording dates. These datasets were consolidated into a master datasheet, allowing records to be organised and filtered by animal ID and date. Animal-level information, including age at calving and breeding values, were stored in a separate table and merged with production records using Animal Birth ID as a unique identifier. Ensuring consistent alignment of repeated production measurements with individual cow attributes.

3.2.2 Analysing Data

Software and Initial Processing

All phenotypic analyses were conducted using SAS software (Version 9.4; SAS Institute Inc). Production records and animal demographic data were imported from Microsoft Excel using PROC IMPORT. Datasets were then merged using Animal Birth ID to ensure correct alignment of repeated production records with individual cow information. Quadratic terms for LWT (weight²) and Days in Milk (DIM²) were generated to allow modelling of non-linear biological relationships commonly observed in lactation curves and body size responses.

Data Cleaning and Quality Control

A structured quality control protocol was applied prior to analysis. Records containing non-biological zero values for key variables: MY, PY, PP, FY, FP and LWT, were excluded, as these were assumed to represent recording or data entry errors. Distributions of continuous variables were assessed using graphical diagnostics, including examination of histograms, density plots, and (Q–Q) plots. Although formal tests of normality were reported, interpretation focused on graphical diagnostics and descriptive statistics due to the large sample size. Linear relationships among traits were assessed using a Pearson correlation matrix generated with PROC CORR.

Descriptive and Phenotypic Regression Analyses

Herd-level temporal trends, represented as changes in milk yield, milk composition traits, and LWT over time, were visualised using daily herd means with 95% confidence limits. Smoothed lactation curves were generated using locally estimated scatterplot smoothing (LOESS) to illustrate production patterns across DIM. To explore maturity effects, cows were categorised into age-at-calving groups representing increasing physiological maturity (2–3 years, 4–5 years, 6–7 years, and ≥8 years). Milk yield and LWT distributions across maturity groups were visualised using boxplots. Individual-level trajectories of LWT and milk yield were plotted across time to confirm the longitudinal structure of the data and to justify the use of repeated-measures statistical models. Linear regression analyses were conducted to quantify phenotypic relationships between LWT and milk production traits, including milk volume, fat percentage, protein percentage, and milksolids yield. For cow i at observation j, the general phenotypic regression model was:

(Eq. 1)

$$Y_{ij} = \beta_0 + \beta_1 Weight_{ij} + \varepsilon_{ij}$$

where Y_{ij} denotes the milk production trait of interest, β_0 is the intercept, β_1 is the regression coefficient for LWT, and ε_{ij} is the residual error term, assumed to be normally distributed with mean zero and variance σ^2. These analyses were performed using PROC REG, and regression coefficients, intercepts, and coefficients of determination (R^2) were extracted and displayed directly on graphical outputs. The purpose of these regressions was to characterise broad phenotypic associations between body size and production traits and to inform subsequent multivariable modelling.

Primary Modelling Framework

To appropriately account for repeated measurements collected from the same animals, linear mixed models were fitted using PROC MIXED. Separate models were specified for milk volume, FP, PP, and milksolids yield. For cow i at observation j, the following linear mixed model was fitted:

(Eq. 2)

$$\begin{aligned} Y_{ij} = \quad & \beta_0 + \beta_1 \,\mathrm{Age}_i + \beta_2 \,\mathrm{DIM}_{ij} + \beta_3 \,\mathrm{DIM}_{ij}^2 + \beta_4 \,\mathrm{Weight}_{ij} + \beta_5 \,\mathrm{Weight}_{ij}^2 \\ & + \beta_6 \,(\mathrm{Age}_i \times \mathrm{DIM}_{ij}) + \varepsilon_{ij} \end{aligned}$$

where Y_{ij} represents the milk production trait of interest, Age_i is age at calving, DIM_{ij} is days in milk, and Weight_{ij} is LWT. Quadratic terms were included to capture non-linear biological responses associated with lactation progression and body size. The interaction between age at calving and DIM was included to test whether lactation trajectories differed across maturity stages.

The repeated animal identifier (Animal Birth ID) was specified as the subject, and within-animal correlation across successive measurements was modelled using a first-order autoregressive [AR(1)] covariance structure. Models were fitted using restricted maximum likelihood (REML).

The AR(1) covariance structure was defined as:

(Eq. 3)

$$\text{Cov}(\varepsilon_{ij}, \varepsilon_{ik}) = \sigma^2 \rho^{|j-k|}$$

where ρ is the autoregressive correlation parameter and $|\, j - k \,|$ represents the temporal distance between observations. This structure assumes that measurements taken closer together in time are more strongly correlated than those further apart, which is biologically appropriate for longitudinal lactation data. Estimated autoregressive correlations were high across all traits, supporting the suitability of this specification. This modelling framework allowed partitioning of production variation attributable to maturity, body size, and stage of lactation while accounting for temporal autocorrelation among repeated measures within cows.

Residual Trait Analysis and Covariance Structure

Residuals from each mixed model were extracted and retained as adjusted phenotypes:

(Eq. 4)

$$r_{ij} = Y_{ij} - \hat{Y}_{ij}$$

where $\hat{Y}_{ij}$ is the model-predicted value for observation j of cow i. These residuals represent variation in production traits after accounting for fixed effects associated with age, LWT, stage of lactation, and longitudinal correlation. Residuals for all traits were merged into a single dataset using animal identifier and herd test date, enabling joint analyses of adjusted phenotypes across traits. Relationships among adjusted phenotypes were evaluated using Pearson correlation analysis conducted with PROC CORR.

Residual-based scatterplots with fitted regression lines were used to visualise the magnitude and direction of residual associations. Regression analysis of adjusted phenotypes was performed using PROC REG. Specifically, residual PP and residual milksolids yield were regressed on residual milk volume to quantify relationships among adjusted traits. Coefficients of determination (R^2) were used to describe the proportion of residual variance explained.

Assessment of Within-Cow Temporal Variation

To characterise within-cow temporal variability in milk yield, descriptive analyses were conducted prior to mixed-effects modelling. Within-cow standard deviations of milk yield were calculated across repeated test-day records for each animal to quantify short-term variability around individual means. Changes in milk yield between successive test days were also computed to examine the magnitude and distribution of temporal fluctuations. These analyses were used to assess whether within-cow variation was random or temporally structured. Residuals from linear mixed models were extracted and treated as adjusted phenotypes, enabling evaluation of temporal dependence and relationships remaining after accounting for age at calving, body size, stage of lactation, and longitudinal correlation. Collectively, these descriptive assessments informed the specification and justification of autoregressive covariance structures in longitudinal models.

Fixed-Effects General Linear Modelling

To complement the phenotypic regressions and to evaluate differences among maturity groups under a fixed-effects framework, a general linear model (GLM) analysis was conducted using PROC GLM in SAS. Milk production traits, including MY, FP, PP, LWT, and MS, were modelled jointly as response variables. Fixed effects included DIM, herd testing date, breed, and categorical age-at-calving group (2–3, 4–5, 6–7, and ≥8 years). Animal Birth ID was included as a random effect to account for repeated observations on the same cows. Least squares means were estimated for age group and breed, and pairwise differences were assessed using standard errors and multiple-comparison procedures. This analysis provided an initial assessment of maturity- and breed-related differences in production traits prior to longitudinal modelling.

3.3 Genetic Analysis using ASReml

In addition to phenotypic modelling, pedigree-based genetic analyses were conducted to estimate additive genetic variation in milk production traits and live weight. Analyses were performed using ASReml, which fits linear mixed models using restricted maximum likelihood (REML) and incorporates pedigree relationships through the numerator relationship matrix. These models partition phenotypic variation into additive genetic, permanent environmental, and residual components, allowing estimation of heritability and repeatability.

3.3.1 Data Preparation and Pedigree Structure

Phenotypic records used for genetic analyses were identical to those described in Sections 3.1-3.2. comprised of repeated test-day observations for MY, FY, PY, MS, FP, PP, and LWT. Data were stored in long format, allowing multiple observations per animal across lactation. Pedigree information was obtained from the farm record systems for the 2022-2023/2023-2024 seasons. Individual pedigree files were cleaned and merged into a single pedigree dataset prior to analysis. The final pedigree file included animal ID, sire ID and Dam ID, unknown parents or missing observations were filled as (".") following ASReml pedigree conventions. The pedigree and phenotypic datasets were match according to animal ID to ensure consistency between records.

3.3.2 Model Specification

Single-trait linear mixed animal models were fitted separately for each production trait. The general model fitted for each trait was as follows.

General statistical model – $Y_{ij} = \mu + X\beta + a_i + p\varepsilon_i + \varepsilon_{ij}$

Where:

Y_{ij} = Vector of phenotypic observations

μ = Overall mean

$X\beta$ = Fixed effects

a_i = Additive genetic (animal) effect

$p\varepsilon_i$ = Permanent environmental effect associated with repeated observations on the same animal;

ε_{ij} = Residual error

Example: Milk yield model - $MY_{ij} = \mu + DIM_i + Breed_i + AG2_k + a_i + p\varepsilon_i + \varepsilon_{ijk}$

Where:

MY_{ij} = Milk yield

μ = Overall mean

DIM_i = Fixed effect for DIM (Days in milk)

$Breed_i$ = Fixed effect for breed

$AG2_k$ = Fixed effect for age

a_i = Additive genetic (animal) effect

$p\varepsilon_i$ = Permanent environmental effect associated with repeated observations on the same animal.

ε_{ij} = Residual error

ASReml was used for all models, and convergence was assessed based on stability of the log-likelihood and variance component estimates across iterations.

3.3.3 Variance Components and Genetic Parameters

For each trait, ASReml was used to estimate genetic variance σ_A^2, permanent environmental variance σ_{PE}^2, and residual variance σ_E^2. Phenotypic variance σ_P^2, was calculated as the sum of these components. Heritability (h^2) was calculated as: $h^2 = \frac{\sigma_A^2}{\sigma_P^2}$

Repeatability (r) was calculated as: $r = \frac{\sigma_A^2 + \sigma_{PE}^2}{\sigma_P^2}$

Variance components, heritability, and repeatability estimates were derived directly within ASReml using VPREDICT statements referencing estimated variance component line numbers.

CHAPTER 4

Results

Dataset Characteristics

After data cleaning and merging, the final dataset comprised 13,489 observations collected from 582 cows, with a maximum of 25 repeated measurements per animal. The dataset included Holstein-Friesian (HF), Jersey (J), and Holstein-Friesian x Jersey (HFxJ) cows. Days in milk ranged from 0 to 199 days (mean = 76.2 days), indicating that observations were concentrated in early lactation but extended into mid-lactation. Mean LWT across all observations was 513.8 kg, average daily milk yield was 20.5 kg, mean daily milksolids production was 1.76 kg, while milk fat and protein percentages averaged 4.60% and 3.99%, respectively. Considerable variation was observed for all traits, with broad ranges reflecting differences among cows, lactation stage, and body size (Table 4.1).

Table 4. 1*:* Summary statistics of the final analytical dataset per measured trait.

Trait	Mean	SD	Minimum	Maximum
Milk volume (L/day)	20.52	5.52	1.10	50.90
Milk fat (%)	4.60	0.73	2.40	11.90
Milk protein (%)	3.99	0.22	3.00	5.50
Milksolids (kg/day)	1.76	0.48	0.15	4.62
Liveweight (kg)	513.83	60.80	302.20	731.00
Days in milk (DIM)	76.23	42.93	0	199

4.1 Descriptive Temporal Trends and Lactation Patterns

Herd-level temporal trends in milk production traits and LWT across the observation period are shown in Figures 4.1 and 4.2. Figure 4.1 presents daily herd means across successive herd test dates, illustrating seasonal patterns in milk yield, milk composition, and LWT at the herd level. Across traits, production changed gradually between test days, with short-term fluctuations occurring around broader temporal trends. When examined across lactation stage, milk yield displayed a clear non-linear relationship with DIM, with an early-lactation increase followed by a gradual decline (Figure 4.2a). This pattern was similarly reflected in milksolids production (Figure 4.2d). In contrast, milk fat and protein percentages showed less monotonic behaviour across DIM (Figures 4.2b and 4.2c), indicating compositional variation was not driven solely by changes in MY. Live weight also varied systematically across DIM (Figure 4.2e), demonstrating coordinated changes in body mass alongside production across lactation stage. Together, these herd-level patterns show that both time and lactation stage contribute substantially to observed variation in milk production traits, motivating further examination of within-cow variation and longitudinal structure in subsequent analyses.

Results

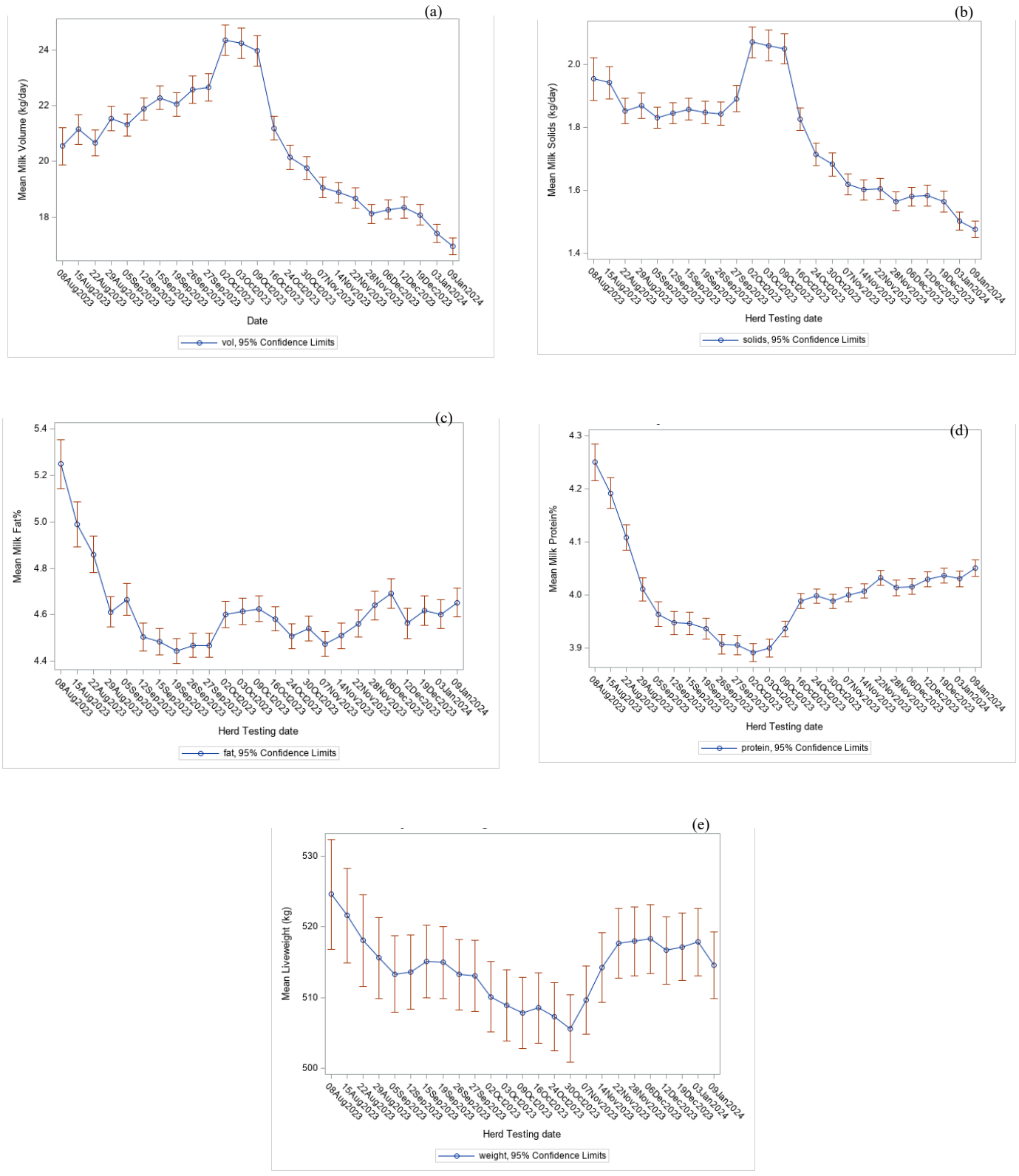


Figure 4. 1: *Herd level mean milk production traits and live weight across successive herd test dates. Points represent herd means for each test day, illustrating temporal patterns in milk yield (a), milksolids (b), milk fat percentage (c) and milk protein percentage (d), and live weight (e) across the observation period.*

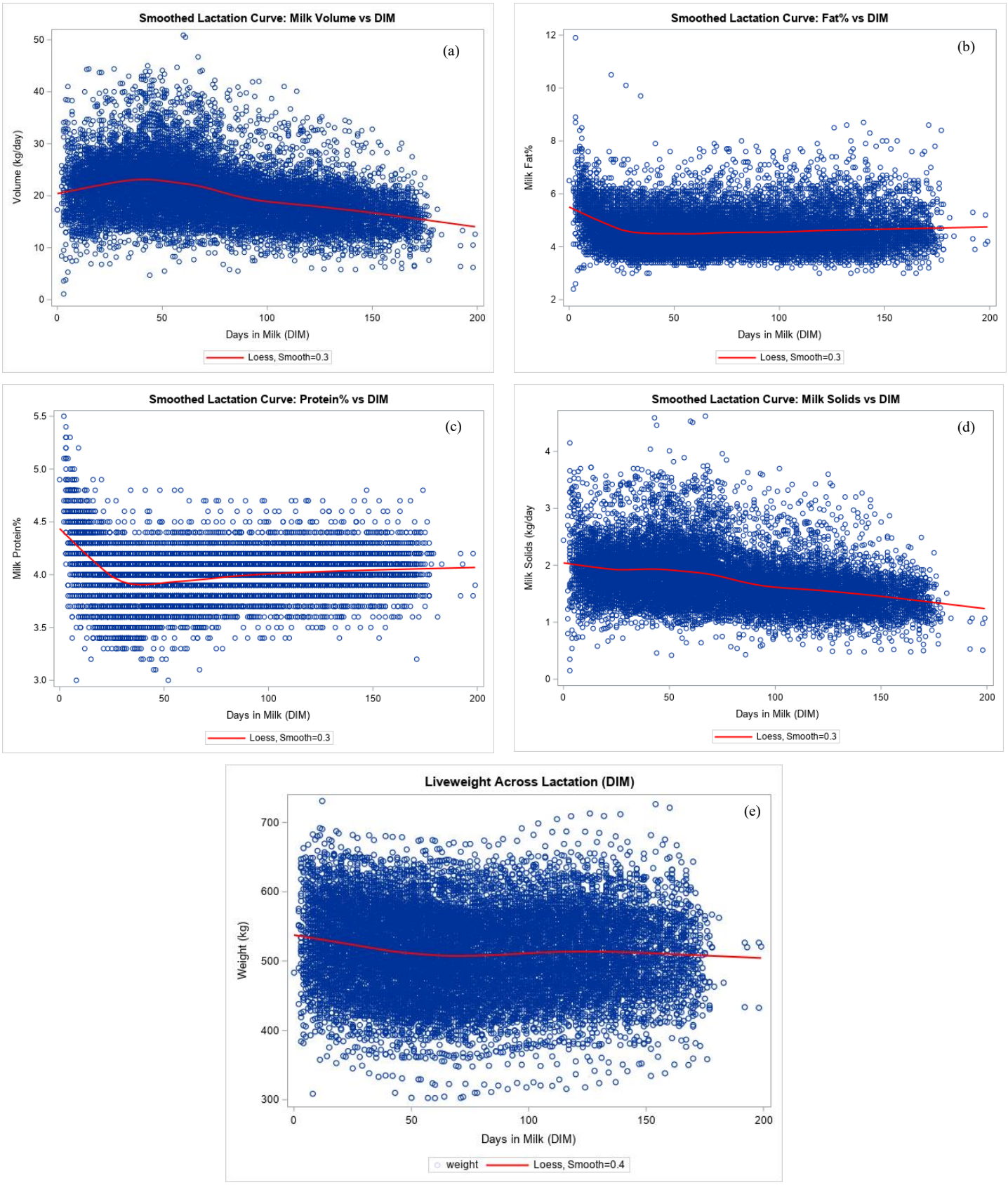


Figure 4. 2: *Herd-level temporal trends in milk production traits and live weight across days in milk (DIM). Milk volume vs DIM (a), milk fat percentage vs DIM (b), milk protein percentage vs DIM (c), milksolids vs DIM (d), and live weight vs DIM (e). Values represent daily herd means with smoothed curves illustrating non-linear lactation patterns for milk yield, milksolids, milk fat and protein percentage.*

4.2.1 Within-Cow Variation and Test-Day Changes in Milk Yield

Within-cow variation in daily milk yield was evident across the study population. The mean within-cow standard deviation of milk yield was 3.30 kg/day, with a median of 2.91 kg/day. Values ranged from 0.77 to 9.39 kg/day, indicating that while some cows maintained relatively stable milk production across test days, others displayed noticeably greater short-term variability. Changes in milk yield between successive herd test days (delta_vol) were generally small and centred close to zero (Figure 4.3). The mean test-day change in milk yield was −0.18 kg/day, with a median change of −0.10 kg/day and a standard deviation of 2.41 kg/day. Approximately 95% of test-day changes lay between −3.2 and +2.7 kg/day. In contrast, a small number of extreme changes were observed, with values ranging from −23.9 to +21.8 kg/day. The distribution of test-day changes showed a strong concentration around zero, with slight negative skewness. This pattern illustrates gradual changes in milk yield between successive recordings rather than abrupt or highly erratic variation across test days. Raw milk yield across days in milk for a subset of cows is shown in Figure 4.4. Individual cows exhibited broadly similar lactation-stage patterns, with milk yield generally declining as days in milk increased, although substantial between-cow differences in absolute production level were evident. These consistent within-cow temporal patterns highlight the strong influence of lactation stage on milk yield and support the inclusion of DIM as a key explanatory variable in the mixed models used to derive adjusted phenotypes.

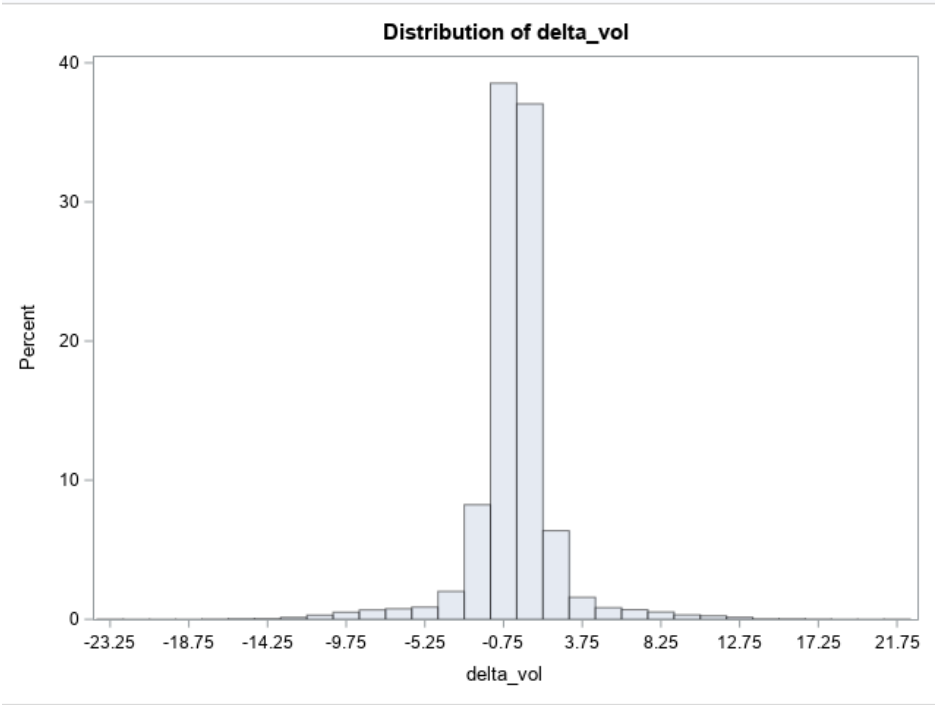


Figure 4. 3*: Distribution of within-cow changes in milk yield between successive herd test days. Delta_vol represents differences in daily milk yield (kg/day) between consecutive test dates for the same cow, illustrating short-term temporal variability in milk production.*

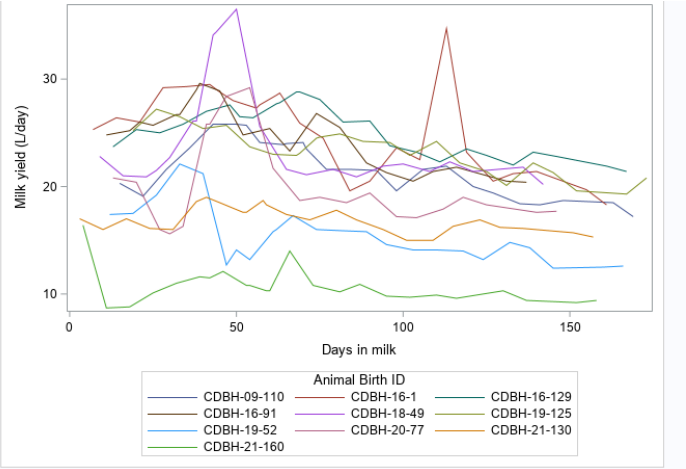


Figure 4. 4*: Daily milk yield across days in milk for individual cows.*

4.2 Phenotypic Associations Between Live weight and Milk Production Traits

Phenotypic associations between LWT and milk production traits were examined using simple linear regression. Live weight was positively associated with milk volume and milksolids, while associations with milk fat and protein percentages were weaker and negative (Figure 4.5).

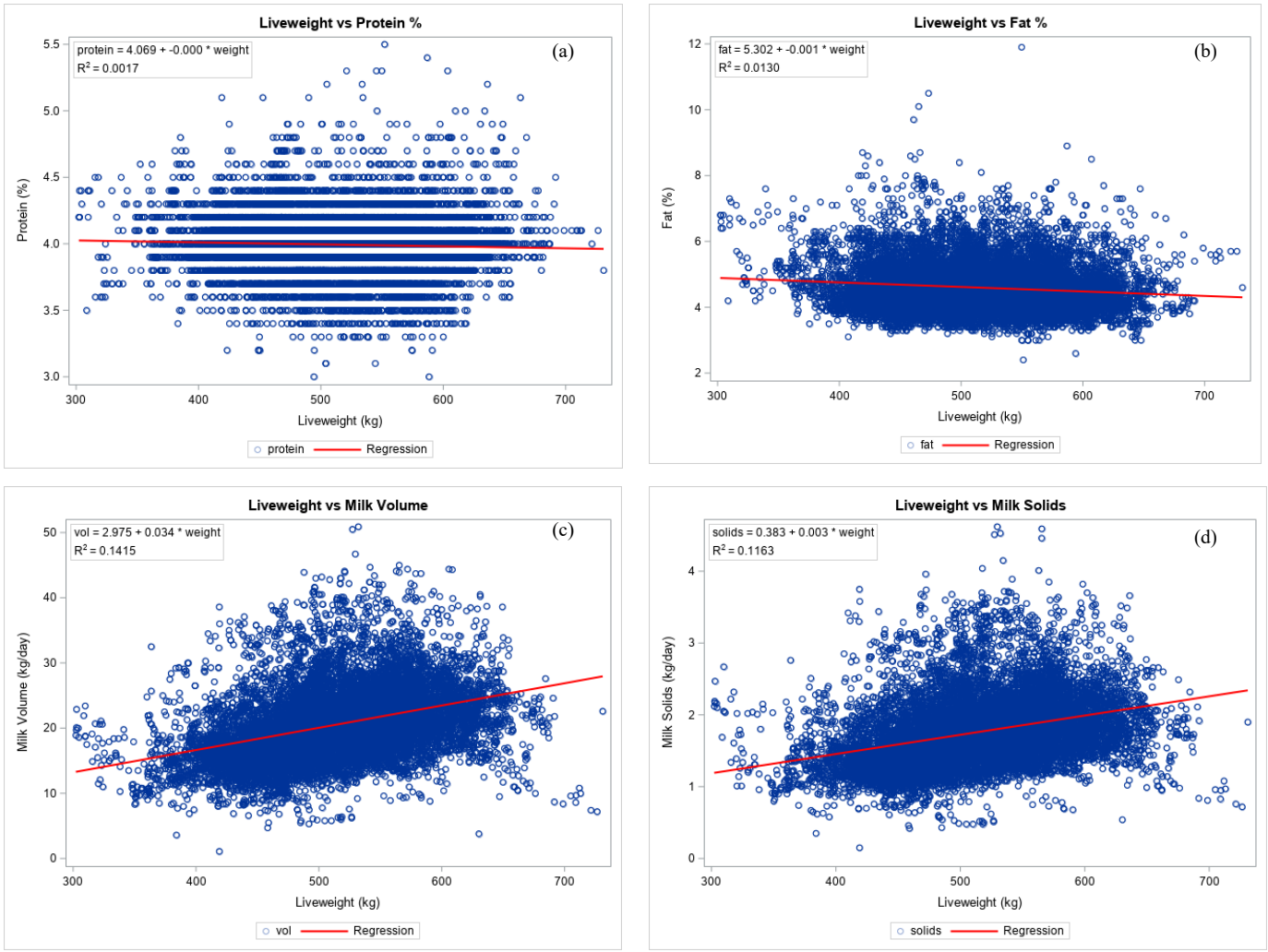


Figure 4. 5*: Phenotypic association between; live weight and milk protein percentage (a), live weight and milk fat percentage (b), live weight and milk volume (c), live weight and milksolids (d). Each point represents a test-day observation, with the fitted linear regression line shown. The coefficient of determination (R^2) is reported on the plot.*

Table 4. 2: Phenotypic regression results for associations between live weight and milk production traits.

Trait	Slope (± SE)	R^2	p-value
Milk volume (kg/day)	0.034 ± 0.00072	0.1415	<0.0001
Milk fat (%)	−0.00137 ± 0.00010	0.0130	<0.0001
Milk protein (%)	−0.00015 ± 0.00003	0.0017	<0.0001
Milksolids (kg/day)	0.00268 ± 0.00006	0.1163	<0.0001

4.3.1 Milk Yield/Volume

A strong positive phenotypic association was observed between LWT and daily MY (Figure 4.5c). Live weight explained 14.15% of the total variation in milk yield ($R^2 = 0.1415$; $p < 0.0001$), representing the strongest associations among the traits observed. Milk yield increased by 0.034 kg/day for each additional kilogram of LWT (Table 4.2).

4.3.2 Milk Fat Percentage

FP showed a weak but significant negative association with LWT (Figure 4.5b). Live weight accounted for 1.30% of the observed variation in FP ($R^2 = 0.0130$; $p < 0.0001$). Fat percentage decreased by 0.00137 percentage units per kilogram increase in LWT (Table 4.2).

4.3.3 Milk Protein Percentage

A similarly weak negative association was detected between LWT and PP (Figure 4.5a). Live weight explained 0.17% of the variation in PP ($R^2 = 0.0017$; $p < 0.0001$), representing the weakest phenotypic association observed. Protein percentage declined by 0.00015 percentage units per kilogram increase in LWT (Table 4.2).

4.3.4 Milksolids

MS was moderately positively associated with LWT (Figure 4.5d). Live weight explained 11.63% of the variation in daily MS production ($R^2 = 0.1163$; $p < 0.0001$). Milksolids increased by 0.00268 kg/day per kilogram increase in LWT (Table 4.2).

Overall, LWT showed stronger phenotypic associations with MY and MS than with milk fat or protein percentages. Regression coefficients and model fit statistics for all traits are summarised in Table 4.2.

4.3 Linear Mixed Model Results

Given the substantial and temporally structured within-cow variation observed in MY (Section 4.2.1), linear mixed models were fitted to account for repeated test-day measurements within cows. Fixed-effect estimates (± SE) and autoregressive correlations from the linear mixed models are presented in Table 4.3. Separate models were fitted for MY, FP, PP, and MS. All models included age at calving, DIM, DIM^2, LWT, LWT^2, and the interaction between age at calving and DIM as fixed effects. A first-order autoregressive [AR(1)] covariance structure was specified to model within-cow temporal correlation. All models converged successfully under the specified covariance structure. Estimated autoregressive correlations indicated strong within-cow temporal dependence across traits, with values ranging from $\rho = 0.846$ to $\rho = 0.930$, confirming that successive test-day records were highly correlated.

Table 4. 3: Fixed effect estimates (± SE) and autoregressive correlations from linear mixed models for milk production traits.

Fixed effect	Milk volume	Milk fat (%)	Milk protein (%)	Milksolids
Intercept	−2.05 ± 4.48	10.11 ± 0.70	4.96 ± 0.21	0.71 ± 0.39
Age at calving	0.97 ± 0.08^{c}	0.057 ± 0.013^{c}	0.011 ± 0.004^{a}	0.099 ± 0.007^{c}
DIM	0.0209 ± 0.0064^{b}	−0.0208 ± 0.0010^{c}	−0.0067 ± 0.0003^{c}	−0.0036 ± 0.0006^{c}
DIM²	−0.00027 ± 0.00003^{c}	0.000096 ± 0.000004^{c}	0.000036 ± 0.000001^{c}	0.000004 ± 0.000003^{ns}
LWT	0.078 ± 0.017^{c}	−0.0167 ± 0.0027^{c}	−0.0037 ± 0.0008^{c}	0.0039 ± 0.0015^{b}
LWT²	−0.00008 ± 0.00002^{c}	0.000014 ± 0.000003^{c}	0.000004 ± 0.000000^{c}	−0.000004 ± 0.000001^{b}
Age × DIM	−0.00250 ± 0.00073^{c}	0.00024 ± 0.00012^{a}	−0.00017 ± 0.00004^{c}	−0.00026 ± 0.00006^{c}
AR(1) (ρ)	**0.865**	**0.901**	**0.930**	**0.846**

a =p < 0.05; b = p < 0.01; c = p < 0.001; ns = not significant (p ≥ 0.05).

4.4.1 Milk Yield/Volume

Milk yield was significantly affected by age at calving, DIM, DIM², LWT, LWT², and the age at calving × DIM interaction (all $p < 0.01$; Table 4.3). Quadratic terms for both DIM and LWT indicated non-linear associations with milk yield across lactation stage and body size. Strong within-cow temporal dependence was observed ($\rho = 0.865$).

4.4.2 Milk Fat Percentage

Fat percentage was significantly affected by age at calving, DIM, LWT, LWT², and the age at calving × DIM interaction ($p < 0.05$; Table 4.3). The quadratic effect of DIM was not statistically significant. Temporal dependence remained strong ($\rho = 0.901$).

4.4.3 Milk Protein Percentage

Protein percentage showed strong associations with lactation stage, with both DIM and DIM² highly significant ($p < 0.0001$; Table 4.3). Age at calving, LWT, LWT², and the age at calving × DIM interaction were also significant. Protein percentage exhibited the strongest within-cow temporal dependence of all traits analysed ($\rho = 0.930$).

4.4.4 Milksolids

Milksolids was significantly affected by age at calving, DIM, LWT, LWT², and the age at calving × DIM interaction ($p < 0.05$; Table 4.3). The quadratic effect of DIM was not significant. However, within-cow temporal dependence remained strong ($\rho = 0.846$).

Across all traits, non-linear effects of predictors and strong within-cow temporal dependence were consistently observed.

4.5 Adjusted Phenotypes and Trait Relationships

Milk production traits were adjusted using the mixed models described earlier to account for the effects of LWT, days in milk, age at calving and herd test date. The residuals from these models were used as adjusted phenotypes and represent variation in milk production that is independent of these systematic effects. Pearson correlations among adjusted MY, FP, PP and MS are shown in Table 4.4. These correlations represent relationships after accounting for the main explanatory variables. The distribution of residuals from the linear mixed models for MY (Figure 4.6) indicates that model assumptions were reasonably well met. Residuals were centred close to zero, therefore the model effectively accounted for systematic variation associated with LWT, DIM, and age. Adjusted MY was strongly positively correlated with adjusted MS ($r = 0.898$, $p < 0.001$), reflecting the contribution of MY to total solids production. The strong positive relationship between adjusted MY and adjusted MS is illustrated in Figure 4.7. In contrast, adjusted MY showed negative correlations with both adjusted PP ($r = -0.348$, $p < 0.001$) and adjusted fat percentage ($r = -0.125$, $p < 0.001$). This indicates that increases in residual MY were generally associated with lower residual concentrations of milk components. Adjusted fat and protein percentages were positively correlated ($r = 0.230$, $p < 0.001$), indicating that cows with elevated residual fat concentration also tended to have elevated residual protein concentration. Adjusted FP was also positively correlated with adjusted MS ($r = 0.249$, $p < 0.001$). In contrast, adjusted PP was weakly negatively correlated with adjusted MS ($r = -0.130$, $p < 0.001$). Pairwise relationships among milk production traits are illustrated in Figure 4.8, which visually summarises the correlation structure described numerically in Table 4.4.

Overall, these results show clear relationships remain among milk production traits even after adjustment for LWT and lactation effects, particularly the strong association between MY and MS and the negative relationships between MY and component concentrations.

Table 4. 4: Pearson correlations among adjusted milk production phenotypes.

Trait	**R_volume**	**R_fat**	**R_protein**	**R_solids**
R_volume	1.000	−0.125***	−0.348***	0.898***
R_fat	−0.125***	1.000	0.230***	0.249***
R_protein	−0.348***	0.230***	1.000	−0.130***
R_solids	0.898***	0.249***	−0.130***	1.000

****p<0.001.*

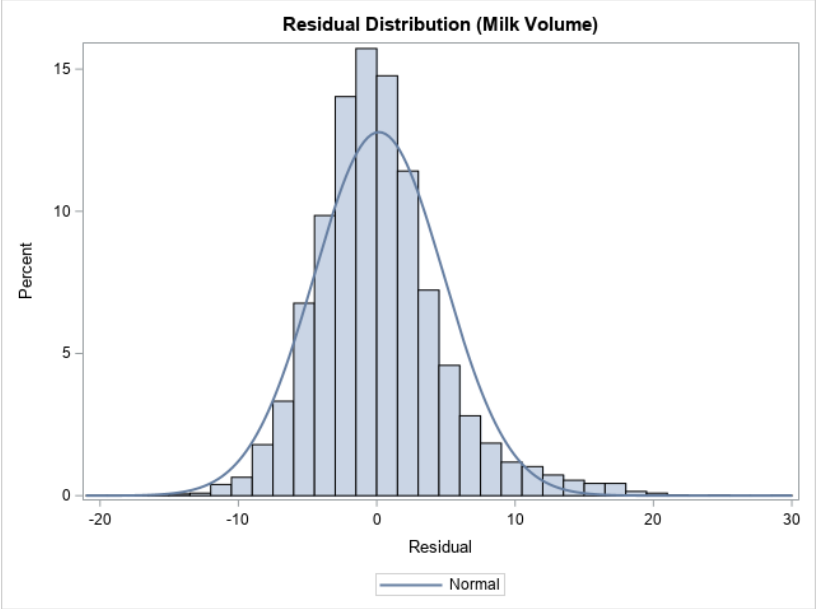


Figure 4. 6: *Histogram of residuals from the linear mixed model for milk volume. Residuals are centred near zero and exhibit moderate right skewness with slightly heavier tails than expected under a normal distribution.*

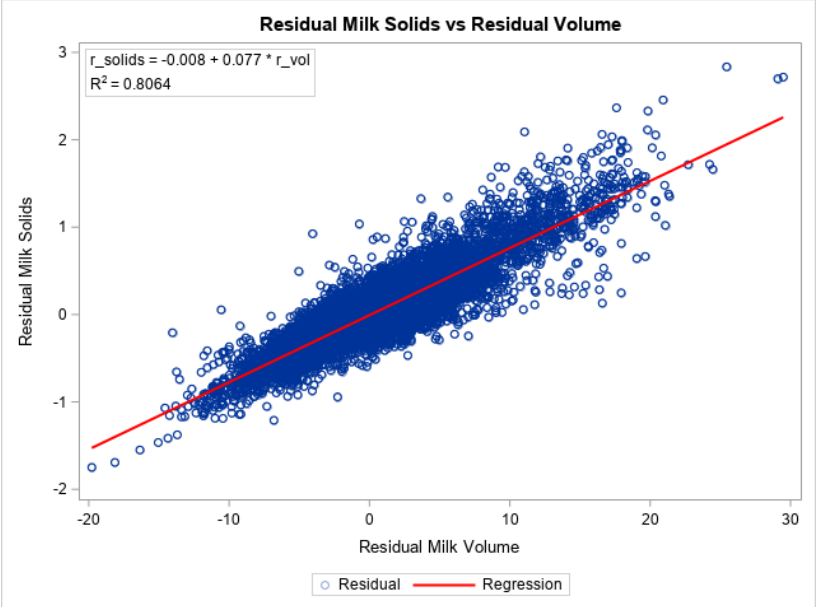


Figure 4. 7: *Relationship between adjusted milk volume and adjusted milk solids.*

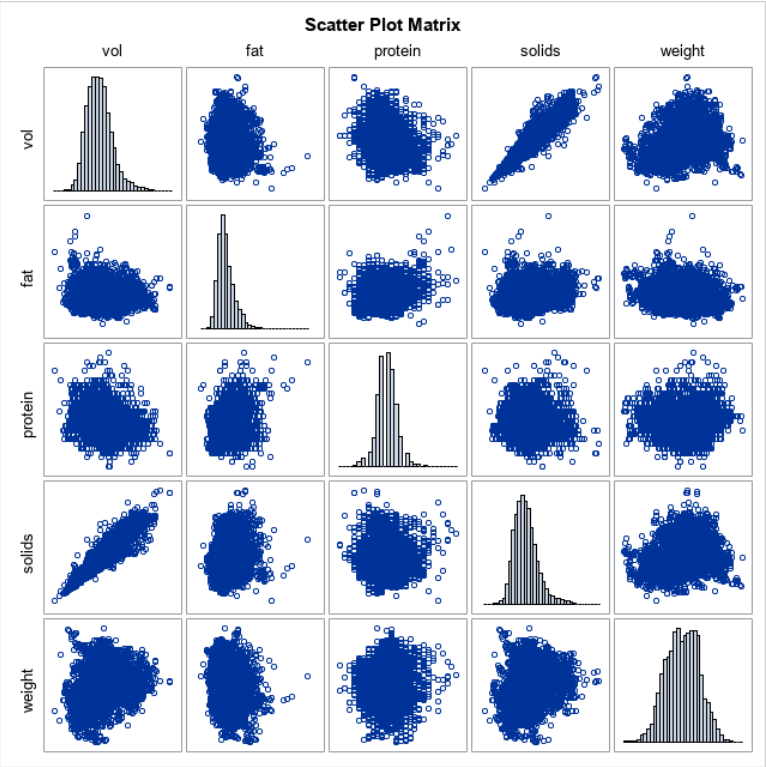


Figure 4. 8: *Scatter plot matrix of adjusted milk production traits (milk volume, fat percentage, protein percentage, and milk solids production).*

4.6 Regression Relationships between Adjusted Phenotypes

The correlation structure among adjusted milk production traits is presented in Section 4.5. To further quantify the strength of these relationships, linear regression analyses were conducted between selected adjusted traits. Regression of adjusted MS on adjusted MY demonstrated a strong positive linear relationship (Figure 4.7). Adjusted MY explained 80.6% of the residual variation in MS ($R^2 = 0.8064$; $p < 0.0001$). This indicates that, after removal of major systematic effects, residual variation in MS production was closely aligned with residual variation in MY. The overall correlation structure among adjusted traits is illustrated in Figure 4.8, with strong positive associations between MY and MS and weaker relationships among component traits. Regression of adjusted PP on adjusted MY identified a weaker but statistically significant negative association (Figure 4.9). Adjusted MY accounted for 12.1% of the residual variation in PP ($R^2 = 0.1209$; $p < 0.0001$), indicating that greater residual MY was generally associated with lower residual protein concentrations. Residual associations between adjusted MY and adjusted FP were weak and explained little residual variation ($R^2<0.05$; $p < 0.001$; results not shown). Given their limited explanatory power, these relationships were not examined further. Evaluation of regression diagnostics indicated no major violations of model assumptions. Residuals were centred close to zero and showed no evidence of strong non-linearity or influential outliers, supporting the use of linear regression models for describing relationships among adjusted phenotypes.

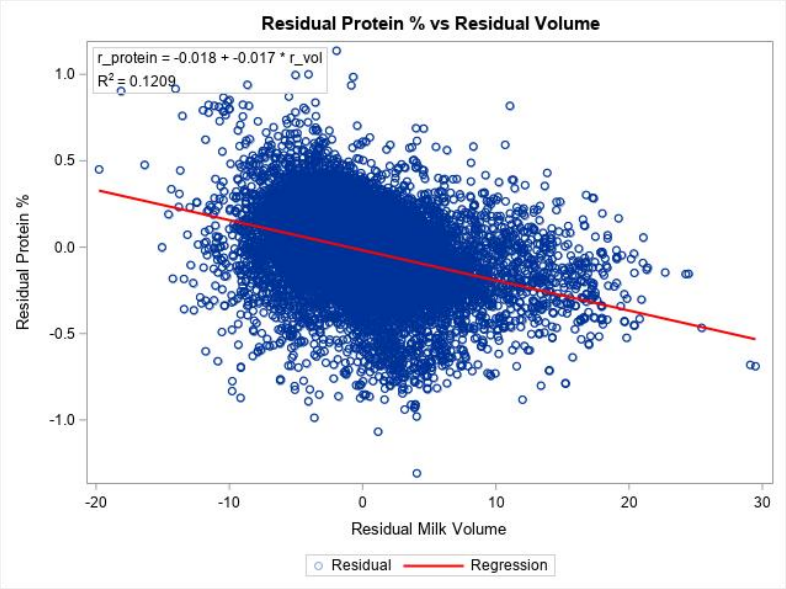


Figure 4. 9: *Relationship between adjusted milk volume and adjusted milk protein percentage following adjustment for fixed effects. The fitted regression line illustrates the weak but significant negative association between the two traits.*

Overall, regression analyses of adjusted phenotypes showed that residual variation in MS was strongly associated with residual variation in MY, whereas relationships between MY and component concentrations were weaker. These patterns indicate that MY remained the primary contributor to variation in MS after accounting for LWT, lactation stage, maturity, and test-day effects.

4.7 Fixed-Effects General Linear Model Results

To complement the phenotypic regression and longitudinal mixed-model analyses, fixed-effects general linear models were used to examine differences in milk production traits across age-at-calving groups and breed categories. Models were fitted using PROC GLM, with DIM included as a covariate and herd testing date, breed, and age-at-calving group treated as fixed effects. Animal Birth ID was specified as a random effect to account for repeated observations. Least-squares means were estimated for age group and breed, with pairwise comparisons used to assess differences among categories.

4.7.1 Main Effects of Age Group and Breed

Significant differences among age-at-calving groups were detected for MY, MS, FP, PP, and LWT (Table 4.5; $P < 0.05$). Least-squares means indicated clear increases in both MY and LWT from first-lactation cows (2–3 years) to mature cows, with values plateauing in older age groups. Adjusted MY increased from ~16 L/day in 2–3-year-old cows to approximately 21 L/day in cows aged 4–7 years,

with no further increase observed in cows aged ≥8 years. Live weight followed a similar pattern, increasing from 451 kg in first-lactation cows to approximately 530 kg in cows aged ≥6 years.

In addition to differences in adjusted means, the distribution of MY differed systematically across age-at-calving groups (Figure 4.10a). Box-and-whisker plots showed that first-lactation cows exhibited lower medians and narrower interquartile ranges for MY compared with older cows. In contrast, mature cows displayed elevated central values and greater dispersion, reflecting increased variability in production among older animals. Upper ranges of MY increased in cows aged ≥4 years, whereas low-volume observations were more common in the youngest age group. A similar pattern was observed for LWT distributions across age-at-calving groups (Figure 4.10b). First-lactation cows showed lower median LWT and a relatively narrow interquartile range, while older cows exhibited progressively increased median LWT and spread. The distribution of LWT appeared to stabilise in cows aged ≥6 years, with comparable medians and ranges observed between the 6–7-year and ≥8-year groups. Milksolids increased progressively with advancing age group, rising from 1.49 kg/day in first-lactation cows to 1.95 kg/day in cows aged ≥6 years. Fat percentage increased modestly across age groups, while PP showed relatively small absolute differences, with first-lactation and oldest cows exhibiting marginally greater values than intermediate age groups.

Table 4. 5: Least-squares means (± SE) for production traits by age-at-calving group, adjusted for DIM, herd testing date, and breed.

Trait	2–3 yrs	4–5 yrs	6–7 yrs	8+ yrs
Milk volume (L/day)	16.3 ± 0.14^{a}	21.0 ± 0.14^{b}	21.5 ± 0.15^{c}	21.1 ± 0.15^{bc}
Milksolids (kg/day)	1.49 ± 0.01^{a}	1.85 ± 0.01^{b}	1.93 ± 0.02^{c}	1.95 ± 0.02^{c}
Milk fat (%)	5.02 ± 0.02^{a}	5.12 ± 0.02^{b}	5.18 ± 0.03^{c}	5.21 ± 0.03^{c}
Milk protein (%)	4.10 ± 0.01^{a}	4.08 ± 0.01^{b}	4.07 ± 0.01^{b}	4.10 ± 0.01^{a}
Live weight (kg)	451 ± 2^{a}	507 ± 2^{b}	526 ± 3^{c}	528 ± 3^{c}

^{a,b,c} *Different superscript letters within a row indicate significant differences among age-at-calving groups or breeds ($p < 0.05$).*

Table 4. 6: Least-squares means (± SE) for production traits by breed (Holstein-Friesian (HF), Jersey (J), and Holstein-Friesian Jersey crossbred (HFxJ), adjusted for DIM, herd testing date, and age-at-calving group.

Trait	HF	HF×J	J
Milk volume (L/day)	22.2 ± 0.13^{a}	20.4 ± 0.14^{b}	17.4 ± 0.16^{c}
Milksolids (kg/day)	1.88 ± 0.01^{a}	1.82 ± 0.01^{b}	1.72 ± 0.02^{c}
Milk fat (%)	4.53 ± 0.02^{a}	5.05 ± 0.02^{b}	5.81 ± 0.03^{c}
Milk protein (%)	3.96 ± 0.01^{a}	4.08 ± 0.01^{b}	4.24 ± 0.01^{c}
Live weight (kg)	543 ± 2^{a}	505 ± 2^{b}	463 ± 3^{c}

^{a,b,c} *Different superscript letters within a row indicate significant differences among age-at-calving groups or breeds ($p < 0.05$).*

Breed was a significant fixed effect for all production traits considered (Table 4.6; $p < 0.05$). Least-squares means reflected distinct breed-specific production profiles. Holstein-Friesian cows produced the greatest adjusted MY (22.2 L/day), followed by HFxJ crossbred cows (20.4 L/day), with J

cows producing the lowest volumes (17.4 L/day). In contrast, J cows exhibited the highest FP (5.81%) and PP (4.24 %), while HF cows had the lowest component concentrations. Milksolids differed significantly among breeds, with HF producing the greatest adjusted yield (1.88 kg/day), followed by HFxJ cows (1.82 kg/day) and J (1.72 kg/day). Live weight varied markedly by breed, with HF cows being the heaviest (543 kg), HFxJ cows intermediate (505 kg), and J cows the lightest (463 kg).

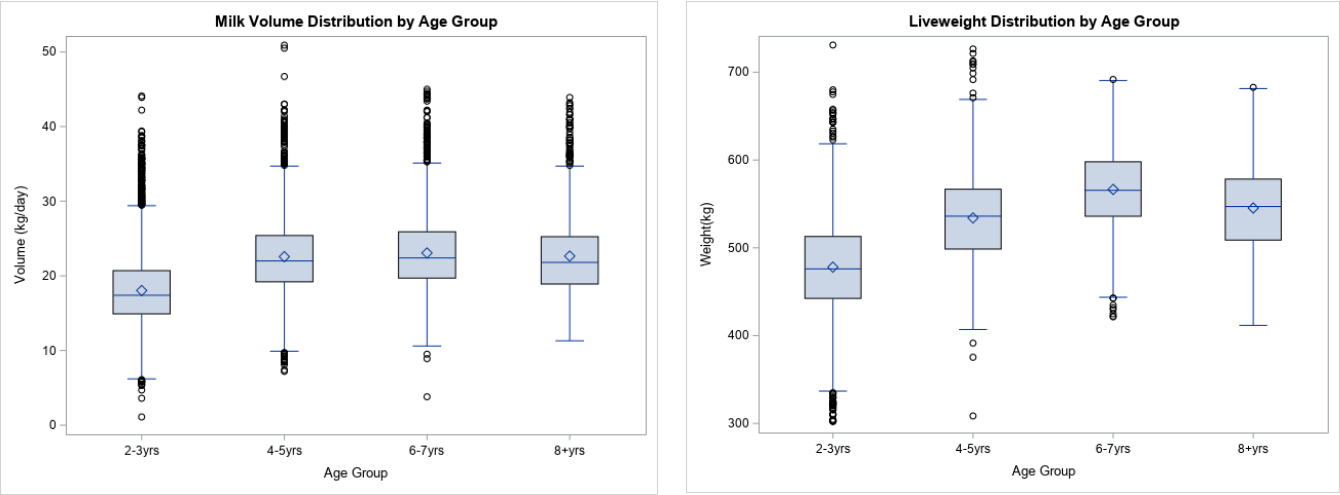


Figure 4. 10*: Distribution of (a) milk volume and (b) live weight across age-at-calving groups (2–3, 4–5, 6–7, and ≥8 years).*

4.7.2 Effect of Days in Milk

Days in milk had a statistically significant effect on all production traits included in the GLM ($p < 0.0001$), indicating consistent changes in milk yield, milk composition, and LWT across lactation stage. These effects remained detectable despite inclusion of DIM as a covariate, reflecting the strong influence of lactation progression on production traits.

4.7.3 Comparison with Mixed-Model Results

Patterns observed in the fixed-effects GLM were consistent with those identified in descriptive and mixed-model analyses presented earlier. Increases in MY, MS, and LWT with advancing age-at-calving group, along with contrasting production profiles among breeds, mirrored trends observed in both raw distributions and adjusted analyses. While the GLM provides clear comparisons among categorical effects, these results are interpreted as complementary to the mixed-effects analyses, which more fully accounted for longitudinal structure and within-cow temporal correlation.

4.8 ASReml Pedigree Results

Variance components, heritability, and repeatability estimates for milk production traits are presented in Table 4.7. These estimates were derived from single-trait animal models fitted using ASReml and reflect

the partitioning of phenotypic variance into additive genetic, permanent environmental, and residual components.

Table 4. 7: Variance components (additive (σ^2A), permanent environment (σ^2PE), residual error (σ^2E), phenotypic (σ^2P)) and genetic parameters (heritability (h2), repeatability (t)) for milk production traits.

Trait	*σ^2A*	*σ^2PE*	*σ^2E*	*σ^2P*	*h^2 (±SE)*	*t (±SE)*
Milk Yield	10.01	19.98	13.66	43.65	0.23 (0.06)	0.69 (0.01)
Fat Yield	0.055	0.095	0.056	0.206	0.27 (0.07)	0.73 (0.01)
Protein Yield	0.013	0.009	0.017	0.039	0.34 (0.06)	0.58 (0.02)
Fat Percentage	0.254	0.247	0.274	0.775	0.33 (0.07)	0.65 (0.01)
Protein Percentage	0.013	0.079	0.048	0.127	0.10 (0.06)	0.62 (0.01)
Milksolids	0.116	0.129	0.111	0.356	0.32 (0.07)	0.69 (0.01)

Heritability estimates varied across traits, ranging from 0.23 for MY to 0.34 for PY, while PP showed a notably lower value (0.10). Milk yield had relatively lower heritability compared to the other traits, indicating a greater influence from environmental or non-genetic factors. In contrast, PY exhibited the highest heritability, indicating a stronger genetic contribution to variation in this trait. Fat yield and milksolids both showed moderate heritability, indicating that these traits are influenced by both genetic and environmental effects. Fat percentage also showed moderate heritability, while PP was substantially lower, signalling limited additive genetic variation for this trait compared with other production traits.

Repeatability estimates were consistently high across all traits, ranging from 0.58 to 0.73. Fat yield showed the highest repeatability, while PY was the lowest. This indicates that a large proportion of variation in production traits is due to permanent environmental effects, and that individual animals tend to maintain their level of performance across repeated records. Overall, these results indicate that most milk production traits have moderate genetic control, while PP is more strongly influenced by non-genetic factors.

CHAPTER 5

Discussion

5.1 Overview of key findings

This study characterised longitudinal patterns of milk production, milk composition, milksolids, and LWT in a New Zealand dairy herd using repeated test-day records. Variation between LWT and milk production presented as highly structured across a lactation rather than random. Although, substantial within-cow variation in milk yield was observed, changes between successive test days were generally small and followed consistent temporal patterns.

Live weight was positively associated with MY and milksolids production, indicating that cattle with larger mass possess greater production capacity. In contrast, relationships between LWT and milk fat or protein percentages were weak, indicating body size contributes little to explaining variation in milk composition. Importantly, these relationships were non-linear, with evidence of diminishing returns as LWT increased. Alluding that while larger body size may support greater output, gains in production do not scale proportionally with increases in maintenance requirements, highlighting an efficiency trade-off with breeding larger cows.

It was evident that strong relationships between milk traits persisted post adjustment for LWT, lactation stage and age. Residual MY remained closely aligned with residual milksolids, whereas negative associations with protein and fat concentration persisted, indicating component dilution reflects intrinsic biological constraints rather than confounding by stage of lactation or body size. Breed and age effects illustrate biological optimisation rather than simply linear improvement. Milk production and liveweight increased alongside maturity before plateauing in older cows. Whereas differences among age at calving groups largely reflected physiological maturity, breed differences reflected contrasting biological and production characteristics. Holstein-Friesian cows produced greater milk volume, whereas Jersey cows produced milk with greater fat and protein concentrations, with crossbred animals generally exhibiting intermediate performance. Combined, these patterns are consistent with the requirements of New Zealand pasture-based dairy systems, which place greater emphasis on milksolids production and biological efficiency than milk volume alone.

5.2 Relationship between Live weight, Milk Volume and Milksolids

As noted in the above overview, LWT was positively associated with both MY and milksolids across the study population showing that larger cows tend to have an improved production capacity. Given that the contrasting negatively associated relationship LWT was seen to have with milk fat and protein percentages, it can be interpreted that body size primarily influences the quantity of produced milk rather than the compositional quality of the milk produced.

This positive association is consistent with previous studies reporting moderate phenotypic correlations between weight and MY (~0.20), indicating trends of larger cows having improved yield capacities

(Sieber et al., 1988; Spelman & Garrick, 1997), with similar relationships observed in New Zealand grazing systems (Alemu et al., 2025). The positive association between LWT and output traits are likely reflecting the increased feed intake capacity, greater visceral mass, and larger mammary tissue volume in heavier individuals. This reflects biological capacity limits, where heavier animals have increased DMI ability, supported by moderate genetic correlations (0.46-0.56) between intake and metabolic body weight (Hardie et al., 2017). However, the positive association between live weight and milk production should not be interpreted as improved feed conversion efficiency. Instead, it reflects the greater biological capacity of larger cows, which require additional feed intake to meet increased maintenance demands.

Importantly, the relationship between LWT and milk production was non-linear, with evidence of diminishing returns as body size increased. This pattern is consistent with Berry et al., (2007), who reported that while milk yield increases with body weight, the marginal response declines as cows become heavier. The significance of quadratic LWT effect indicates that incremental increases in mass did not result in proportional gains in MY or milksolids yield. Sieber et al., (1988) identified negative relationships between body weight and solids-not-fat concentration (−0.14 to −0.20), supporting the concept that increases in MY are often accompanied by dilution of milk components. This illustrates the existence of biological limits beyond which additional size contributes more to maintenance cost than to productive output.

From an efficiency perspective, these findings highlight a fundamental trade-off between production capacity and energetic cost. What larger cows produce in MY they cancel out with increased requirements for basal metabolism which reduces the marginal efficiency of additional body mass (Schmidtmann et al., 2023). Sieber et al., (1988) reported significant negative correlations (−0.33) between body weight and feed efficiency, demonstrating that larger cows do not necessarily convert feed into milk more efficiently. Consequently, selection for increased LWT risks inflating maintenance expenditure without delivering proportional returns in milk-solid production, particularly in systems where intake is constrained. These results are especially relevant within pasture based dairy systems, where feed availability can be limited. Under such conditions, increased body size may emphasise inefficiency by diverting a more energy towards maintenance instead of saleable product. Lembeye et al., (2016) and Prendiville et al. (2009) reported that smaller Jersey cows outperform larger Holstein-Friesians in efficiency when expressed as milksolids per unit of liveweight, particularly under constrained feeding conditions. These observations support breeding strategies that favour a moderate body size that instead focuses on efficient milksolids production and not MY. Spelman and Garrick (1997) demonstrated that incorporating liveweight with a negative economic weighting into selection indices resulted in an additional 3–5% economic gain compared to selection on production traits alone.

5.3 Lactation Stage and Within-Cow Variation

Consistent with the dynamic nature of lactation, there was considerable within-cow variation in milk yield was observed across the study population. Which is consistent with longitudinal studies of milk production, where within-cow variation is structured rather than random due to underlying biological regulation (Krattenmacher et al., 2019). Individual cows exhibited short-term fluctuations in milk production between successive test days, reflecting ongoing physiological and environment influences throughout lactation. However, this variation was not random, Krattenmacher et al (2019) demonstrated that repeatability of energy-corrected milk remains high (>0.78) across lactation, indicating that production at one time point is predictive of production at another.

Changes in milk yield between test days were modest and followed consistent temporal patterns, indicating strong dependence between successive observations. Similarly, genetic correlations between different stages of lactation remain positive and substantial (>0.59), reinforcing the strong dependence between successive observations (Krattenmacher et al., 2019). The high within-cow temporal correlations that were observed demonstrate that production at any given point in lactation is strongly influenced by previous physiological states. This pattern has also been reported in New Zealand populations, where repeatability estimates of 0.40–0.52 indicate consistent performance within cows across the lactation cycle (Lembeye et al., 2016). Sneddon et al., (2015a) also observed high repeatability across test-day production traits, concluding that cows tend to maintain their relative production ranking throughout the season. The underlying biology of lactation is reflected by these patterns. Nutrient partitioning and milk synthesis are regulated in a coordinated manner across time rather than independently responding per measurement. Subsequent research has shown that this process is regulated by endocrine signalling, including growth hormone and insulin, which maintain lactation performance through coordinated metabolic control (Ledinek et al., 2019).

This coordinated regulation has been widely described in dairy cattle, where patterns of production and energy balance reflect a structured biological response rather than independent daily variation (Friggens et al., 2007). This strong temporal continuity observed in milk production aligns well with the foundational concept of homeorhesis, well described by Bauman and Currie (1980). Under this principle, endocrine and metabolic systems coordinate the partitioning of nutrients to support a directed physiological trajectory such as lactation.

There are important implications that arise from recognising these patterns and variation. Treating repeated test-day records as independent observations would ignore the strong temporal dependence inherent in lactation data and risk mischaracterising biological relationships. Including DIM and quadratic lactation terms, alongside an autoregressive covariance structure was thus justified and necessary to model longitudinal production patterns and isolate meaningful residual variation for

subsequent analyses (Friggens et al., 2007). The inclusion of DIM and higher-order terms is consistent with standard modelling approaches, where polynomial functions are used to capture the non-linear trajectory of lactation curves (Dechow et al., 2004; Lembeye et al., 2016). Similarly, random regression models used in genetic evaluations rely on these approaches to accurately partition variance and avoid confounding longitudinal biological patterns with residual variation (Haile-Mariam & Pryce, 2017; Sneddon et al., 2015a).

5.4 Adjusted Trait Relationships and Component Dilution

After adjusting for LWT, age at calving and stage of lactation, strong relationships among the milk traits remained. These persistent relationships are consistent with approaches using adjusted phenotypes or yield deviations, where core trait relationships remain after accounting for environmental and physiological effects (Pryce & Harris, 2006). Berry & McCarthy, (2021) reported very high genetic correlations (≈0.91) between adjusted and unadjusted milksolids traits, indicating that these relationships are intrinsic rather than driven by confounding factors such as body size or maturity. Residual MY was closely associated with residual milksolids, while negative relationships between residual MY and both residual protein and fat concentration persisted. These findings indicate that core trait relationships were not solely driven by differences in body size, maturity or lactation stage. Comparable findings from mixed-model studies indicate that relationships between MY and milksolids persist even after correcting for systematic effects, confirming that they are fundamental to mammary physiology (Lembeye et al., 2016; Schopen et al., 2011).

This strong association between adjusted MY and adjusted milksolids indicates that residual variation in MS production remains strongly driven by differences in MY even after accounting for major systematic effects. Sneddon et al., (2015b) reported phenotypic correlations of 0.92 between milk yield and protein yield in New Zealand dairy cattle, confirming that increases in MY are tightly coupled with increases in total component yield. Sieber et al., 1988 also reported high correlations between MY and fat yield (0.94) and MY and energy-corrected milk (0.98). This reflects the fact that the amount of fat and protein exported in milk is closely linked to the total volume of milk secreted. As a result, residual increases in MY are accompanied by corresponding increases in MS yield despite concurrent changes in component concentration. In contrast to MS yield, the persistence of negative relationships between residual MY and residual protein and fat concentrations indicates that component dilution reflects biological constraints rather than the effect of lactation stage or body size. This dilution effect has been observed, with negative phenotypic correlations reported between MY and fat percentage (−0.13) and solids-not-fat percentage (−0.17) (Sieber et al., 1988). Milk yield is largely regulated through the osmotic effects of lactose synthesis, which initiates the movement of water into mammary secretory vesicles and therefore determines milk volume (Neville, 2001). Consequently, MY is strongly associated

with lactose yield (Lopdell et al., 2017), whereas fat and protein concentrations, may decline as milk volume increases if solids synthesis does not increase proportionally.

These findings reinforce the concept that MY and milk composition are governed by partially independent physiological processes. This is supported by genomic studies identifying antagonistic pleiotropy in key loci such as DGAT1 and ABCG2, which simultaneously influence MY and component concentration in opposing directions (Cohen-Zinder et al., 2005; Littlejohn et al., 2014). While MY can increase via enhanced water secretion, the synthesis of fat and protein is restrained by substrate availability, enzymatic capacity and mammary secretory limits. This clearly illustrates the risk of selection for high MY without ensuring synthesis of milksolids increases proportionally of diluting component concentration. Which emphasises the importance of understanding these milk composition and MY trade-offs when evaluating production efficiency, particularly in systems where payment for product is based on milksolids and not solely fluid volume.

5.5 Age and Breed Effects on Production and Size

Milk production and LWT increased with advancing age at calving, reflecting the biological maturation of cows as they progress from first lactation into full physiological maturity. This pattern reflects well-established lactation biology, where production traits increase with maturity due to continued skeletal growth and increasing metabolic capacity (Friggens et al., 2007; Roche et al., 2007). Cows in their first lactation exhibited lower MY, MS yield and LWT relative to the older cows. This reduced performance in early lactation is widely attributed to competition for nutrients between growth and lactation, with younger cows prioritising body tissue accretion over milk synthesis (Buckley et al., 2000; Dechow et al., 2001). Importantly, increases in milk production and LWT were non-linear across age groups. Production gains plateaued from approximately six years of age onward, with little additional increase observed in older cows. Such plateauing reflects the attainment of a mature physiological state, where cows reach equilibrium between growth, maintenance, and production rather than continued increases in performance (Friggens et al., 2007). This indicates that the attainment of a biological optimum rather than a progressive decline or constraint in older cows. The seemingly lower performance observed in first lactation cows should not be used to interpret their production value as these differences are due to developmental limitations associated with ongoing growth and prioritisation of body tissue accretion over maximal milk production. As cows mature, resources are increasingly allocated toward lactation due to reaching physiological maturity deeming growth no longer a priority (Bauman & Currie, 1980). This results in greater and more stable milk output in subsequent lactations.

There is clear evidence of breed-specific production profiles which illustrated contrasting production strategies. These breed differences are consistently reported in dairy systems, where Holstein-Friesians are selected for high MY while Jerseys are selected for elevated milk component concentrations

(Lembeye et al., 2016; Vance et al., 2013). Holstein-Friesian cows were the heaviest animals and produced the greatest MY however, they exhibited lower milk fat and protein concentrations. In contrast, Jersey cows were lighter and produced lower MYs, but they exhibited substantially greater component concentrations in their milk. The greater component concentration observed in Jerseys is consistent with genetic and physiological differences in milk synthesis pathways, including variation in fat metabolism genes such as AGPAT6 (Littlejohn et al., 2014).

Following typical heterosis, Holstein-Friesian x Jersey crossbreds, displayed performance above the mid-parent average across production traits and body size, reflecting the combined influence of parental breeds (Prendiville et al., 2009; Lembeye et al., 2016; Vance et al., 2013). The production profile represents a biological compromise between volume-driven and concentration-driven strategies and aligns well with pasture-based systems that require robustness and high reproductive performance under elemental conditions. Genotype x environment interactions further reinforce this, with high-production genotypes often performing less efficiently under pasture-limited systems where intake is constrained (Fulkerson et al., 2008). This is particularly important when the New Zealand industry predominantly pays farmers based on milksolids (kgMS) rather than total volume which aligns with the focus on manufacturing concentration commodity products such as milk powder, butter and cheese (Fonterra Co-operative Group Limited, 2025).

This pricing structure is designed to incentivise farmers to produce high quality milk instead of striving for simply increased volumes as lower concentrate milk ends up costing more in transporting, processing and manufacturing without proportional increases in product values. Collectively, the observed age and breed effects therefore support selection strategies focused on balancing energy allocation, production efficiency and long-term sustainability. These selection strategies reflect the underlying polygenic architecture of milk production traits, where both large-effect loci (DGAT1, ABCG2) and many small-effect genes collectively influence productivity and efficiency (Hayes et al., 2010; Jiang et al., 2018; Krattenmacher et al., 2019). In turn, such strategies benefit farm profitability, while also reducing production costs and enhancing product value and efficiency for milk manufacturers.

5.6 Heritability and Repeatability

Heritability estimates from the analysed milk production and LWT traits ranged from low to moderate (0.10 to 0.34), indicating both genetic and environmental factors influence phenotypic variation. Most trait heritabilities fell within a moderate range, PP exhibited the lowest heritability at 0.10 whereas PY which had the highest estimate of 0.34. Milk yield exhibited a heritability of 0.23, FY was 0.27, FP was 0.33 and MS was 0.32. These estimates are generally consistent with previous studies in both New Zealand populations (Jayawardana et al., 2021; Lembeye et al., 2016; Sneddon et al., 2015a) and overseas dairy populations (Schmidtmann et al., 2023; Fang et al., 2017). In New Zealand cattle,

Sneddon et al., (2015a) reported heritabilities of 0.22 for MY, 0.19 for FY, 0.16 for PY, 0.35 for FP and 0.32 for PP, while Lembeye et al., (2016) reported comparable estimates ranging from 0.21–0.26 for FY and 0.22–0.25 for PY. Jayawardana et al., (2023) similarly reported heritabilities of 0.23 for MY, 0.18 for FY, 0.13 for PY, 0.63 for FP and 0.59 for PP. In overseas populations, Schmidtmann et al., (2023) reported heritabilities of 0.33, 0.28 and 0.24 for MY, FY and PY respectively, while Fang et al., (2017) reported estimates approaching 0.39 for both FY and PY in Holstein cattle. Overall, the majority of estimates from this study fall within the range commonly reported for milk production traits and indicate sufficient additive genetic variation to enable continued response to selection.

An interesting result from the present study was the contrast between the heritability estimates for PY (0.34) and PP (0.10). While the estimate for PY was comparable to, or greater than, those reported in previous studies, the PP estimate was substantially lower than values commonly reported in both New Zealand and international dairy populations. PP heritability estimates of 0.32, 0.37, 0.59, and 0.67 have been reported by Sneddon et al., (2015a), Lopez-Villalobos et al., (2020), Jayawardana et al., (2023), and Lembeye et al., (2016), respectively. Even larger estimates have been reported by Schopen et al., (2011) in Dutch Holsteins (0.66). The comparatively low estimate observed in this study suggests that a greater proportion of the phenotypic variation in PP may have been attributable to environmental influences, measurement error, or other non-additive sources of variation. It may also indicate that variation in protein concentration was more strongly influenced by changes in MY and the associated dilution effect than by additive genetic differences among animals, as increases in milk volume can reduce component concentrations unless solids synthesis increases proportionally (Haile-Mariam & Pryce, 2017; Neville, 2001). Given the consistency of higher PP estimates reported across numerous studies, this result should be interpreted cautiously and may warrant further investigation when using larger populations and additional seasons of data.

The magnitude of these estimates is also consistent with expectations for polygenic quantitative traits, where variation is controlled by many loci of small effect (Hayes et al., 2010; Hayes & Daetwyler, 2019). However, the low heritability estimate for PP suggests that a substantial proportion of phenotypic variance was not explained by additive genetic effects within the animal model. Several factors may have contributed to this result, including biological constraints governing protein synthesis, limitations associated with the accuracy of sensor-derived protein estimates, and characteristics specific to the study population. In addition, increases in MY are often accompanied by dilution of milk component concentrations unless solids synthesis increases proportionally (Neville, 2001; Haile-Mariam & Pryce, 2017). Consequently, variation in PP may be more strongly influenced by MY dynamics than by additive genetic differences alone.

Methodological differences among studies may explain why the PP heritability estimate obtained in the present study was lower than those reported previously. The present study utilised repeated test-day

records and included permanent environmental effects to account for repeated observations on the same animal. Whilst this approach more accurately partitions sources of variation within longitudinal data, it can also reduce the proportion of variance attributed to additive genetics and therefore result in lower heritability estimates than models which do not explicitly separate permanent environmental effects. Schmidtmann et al., (2023) similarly demonstrated that estimates of genetic parameters can vary depending on population structure, trait definition, and model specification, despite analysing comparable milk production traits. Consequently, heritability estimates are inherently population- and model-specific, and differences among studies are expected due to variation in breed composition, management system, statistical methodology and environmental conditions (Jayawardana et al., 2023; Lembeye et al., 2016).

Repeatability estimates were consistently moderate to high across all traits, ranging from 0.58 to 0.73. Fat yield exhibited the greatest repeatability (0.73), while PY exhibited the lowest (0.58). These values are comparable with previous studies reporting repeatability of approximately 0.40–0.52 for test-day production traits in New Zealand dairy populations (Lembeye et al., 2016) and greater than 0.70 for some milk production traits when repeated records across lactation are considered (Krattenmacher et al., 2019). Similarly, Sneddon et al., (2015a) reported strong repeatability across test-day milk traits, suggesting that cows generally maintain their relative performance ranking throughout the lactation. The moderate to high repeatability estimates observed in the present study therefore indicate that a large proportion of variation in milk production traits was attributable to permanent animal effects rather than temporary environmental influences. This suggests that cows generally maintain relatively consistent production rankings throughout lactation, supporting the use of early records as indicators of future performance and reinforcing the value of repeated measurements in genetic evaluation.

Trait-specific differences in heritability provide further insight into the genetic architecture of milk production traits. Yield traits exhibited moderate heritability, confirming their capacity to respond to selection. Fat percentage and MS production similarly demonstrated moderate genetic control, consistent with previous studies reporting moderate to high heritability milk composition traits (Jayawardana et al., 2023; Sneddon et al., 2015a). Established genetic relationships among milk production traits indicate strong positive genetic correlations among yield traits, resulting in favourable correlated responses to selection, whereas negative relationships between MY and milk component concentrations are commonly associated with the dilution effect (Haile-Mariam & Pryce, 2017). Consequently, selection focused solely on increasing MY may reduce milk component concentrations unless solids synthesis increases proportionally, highlighting the importance of balanced, multi-trait breeding objectives that simultaneously consider volume, composition and overall production efficiency (DairyNZ & LIC, 2025; Neville, 2001).

5.7 Methodological Considerations and Model Performance

A key strength of this study was the use of repeated test-day records collected from the same cows across a lactation. This longitudinal structure enabled explicit modelling of within-cow temporal variation, rather than relying on single measurements or lactation averages that could potentially obscure biologically meaningful dynamics (Haile-Mariam & Pryce, 2017; Krattenmacher et al., 2019; Sneddon et al., 2015a). Friggens et al (2007) emphasised that lactation is a coordinated biological process, meaning that simplified summary measures may fail to capture underlying adaptive strategies in nutrient partitioning. Repeated measurements increase statistical power and enable the separation of temporary environmental effects from persistent production patterns across lactation, thereby strengthening inference regarding relationships among milk production traits and LWT (Haile-Mariam & Pryce, 2017; Lembeye et al., 2016; Sneddon et al., 2015a). This increased resolution has been shown to improve estimation of production trends and trait relationships, as large datasets of repeated measures reduce noise and enhance statistical power relative to single observations (Auldist et al., 2007; Walsh et al., 2008). The specification of a first-order autoregressive (AR(1)) covariance structure was biologically and statistically appropriate for this type of data. The use of AR(1) structures is well established in dairy science for modelling repeated measures, as they effectively capture the declining correlation between observations as the time interval increases (Walsh et al., 2008).

The consistently high autoregressive correlations across all traits confirmed that observations closer together in time were more strongly related than those further apart, aligning with the expected continuity of lactation processes. This temporal dependence reflects biological continuity in lactation, where production is regulated through coordinated endocrine and metabolic processes rather than independent daily responses (Friggens et al., 2007). This approach allowed the models to account for temporal dependence while maintaining a relatively simple and interpretable covariance structure. More complex structures could have been explored; however, the strong and consistent AR(1) correlations indicate that this specification was sufficient to capture the key longitudinal patterns present in the data without unnecessary overparameterisation. Alternative covariance structures, including unstructured, compound symmetry, and heterogeneous autoregressive models, can also be applied to repeated-measures data. However, AR(1) provided an appropriate representation of the correlation structure in the present study while maintaining computational simplicity and interpretability (Walsh et al., 2008; Pires et al., 2013).

Another important methodological decision was the use of residuals from the mixed models as adjusted phenotypes. This approach mirrors the concept of “yield deviations,” which are widely used in genetic evaluations to remove systematic environmental effects and isolate intrinsic variation in production traits (Pryce & Harris, 2006). By removing the effects of LWT, age at calving, stage of lactation, and herd test date, the residuals represented variation in production traits that could not be attributed to these known

systematic influences. Similar pre-adjustment procedures have been used in genomic and phenotypic studies to ensure that observed relationships reflect underlying biology rather than confounding environmental influences (Schmidtmann et al., 2023; Schopen et al., 2011). This allowed clearer interpretation of intrinsic relationships among milk production traits, particularly when examining component dilution and trade-offs. Berry and McCarthy (2021) similarly used phenotypic residuals to examine trait relationships independent of body size and lactation stage, demonstrating the effectiveness of this approach in isolating biological constraints. The persistence of strong relationships among adjusted traits, especially the close alignment between residual MY and MS, demonstrates that these associations reflect underlying biological mechanisms rather than confounding factors. The persistence of these relationships after adjustment supports findings that core associations among milk traits are driven by inherent biological mechanisms, such as the dilution effect and shared metabolic pathways (Lopdell et al., 2017).

Using adjusted phenotypes also provided a bridge between phenotypic and genetic interpretation. While mixed models accounted for major environmental and physiological effects, residual variation still contains both genetic and unexplained environmental components. This is consistent with genomic studies that utilise pre-corrected phenotypes to identify pleiotropic loci influencing multiple production traits (Jiang et al., 2018). As a result, relationships observed among adjusted traits can be interpreted as more directly reflecting inherent biological constraints and trade-offs. This approach is particularly valuable in systems like in New Zealand, where environmental variation across lactation can otherwise mask important underlying relationships (Pryce & Harris, 2006).

5.8 Implication and Practical Applications of Findings

The findings of this study have several important implications for both dairy breeding programmes and on-farm decision-making within pasture-based systems. At a fundamental level, the results reinforce that milk production traits cannot be considered in isolation, due to being tightly linked through biological processes governing energy allocation, metabolic balance, and mammary function (Bauman & Currie, 1980; Friggens et al., 2007). In particular, the persistence of negative relationships between MY and component concentrations (PP and FP) highlight the importance of avoiding selection strategies that prioritise volume alone. This antagonistic relationship has been widely reported, with increases in milk yield often associated with reductions in fat and protein concentration due to biological dilution effects (Fulkerson et al., 2008; Haile-Mariam & Pryce, 2017; Sieber et al., 1988). At the genetic level, this trade-off is reinforced by pleiotropic loci such as DGAT1 and ABCG2, where alleles that increase component concentration frequently reduce total MY (Cohen-Zinder et al., 2005; Jiang et al., 2018).

Within the New Zealand context, these findings strongly support the structure of the BW index, which places emphasis on milksolids production rather than fluid volume (DairyNZ, 2025; Spelman & Garrick,

1997). The observed dilution effect confirms that increases in MY without proportional increases in fat and protein synthesis result in reduced product value and processing efficiency (Lopdell et al., 2017). As a result, selection decisions that target milksolids yield rather than raw volume are better aligned with both biological constraints and economic outcomes (Auldist et al., 2007).

The relationship between LWT and production traits emphasis the centrality of efficiency-based selection. As demonstrated, larger cows did show an increased production capacity, but with a non-linear relationship the return diminishes at greater body sizes (Hardie et al., 2017; Sieber et al., 1988). This is consistent with Prendiville et al. (2009), who reported that Jersey and HF×J cows achieved greater milk-solids production relative to body weight and energy intake than Holstein-Friesians in pasture-based systems. Berry et al (2007) similarly reported diminishing returns in milk production as body weight increases, indicating that marginal gains decline as LWT increases. Thus, selecting for increased LWT in unlikely to result in proportional gains in MS production and instead increases maintenance requirements. In New Zealand's pasture-based systems where feed is constrained and has a milksolids driven dairy payment system, it is clearly an inefficient selection strategy.

From a practical perspective, the increasingly available automated technologies such as walk-over weighing and in-shed milk recording systems offer opportunities to integrate these findings into on-farm management. High-frequency data collection allows farmers to monitor changes in LWT and milk production in real time (Alawnah et al., 2012; Alemu et al., 2025; Krattenmacher et al., 2019; Sneddon et al., 2015a), enabling more informed decision-making regarding feeding strategies, mating selection and culling. The ability to track within-cow trends across lactation provides insight into how individual animals are performing and responding to changes in lactation stage, nutritional status, and management conditions, thereby supporting targeted management interventions.

These results also indicate that a focus on maintaining adequate BCS across lactation is essential for reproductive performance and long-term-performance. This is consistent with the review of Pryce and Harris (2006), which reported favourable genetic relationships between BCS and fertility traits in New Zealand dairy cattle, indicating that cows maintaining body reserves are more reproductively successful. The observed relationships between production, LWT and milk component traits further highlight the key role of balancing energy allocation rather than maximising output in a single dimension. This balance is critical, as excessive mobilisation of body reserves during early lactation is associated with increased risk of metabolic disorders and reduced long-term MS production, fertility, and longevity (Roche et al., 2009; Sun et al., 2025). Breeding strategies that incorporate and prioritise BCS and health traits alongside production are better positioned to produce animals both resilient and efficient in New Zealand dairy systems.

5.9 Limitations of the Study

While this study provides valuable insight into the relationships among LWT and milk production traits, several limitations should be considered when interpreting the results. The most significant limitation is the use of data from a single herd from one farm. This limitation is well recognised in dairy research, as herd-specific management and environmental conditions can significantly influence production traits and limit broader applicability of findings (Pryce & Harris, 2006). Although the dataset was detailed and included repeated measurements across lactation, it will not fully represent the range of genetic diversity, management practices or environmental conditions present across the wider New Zealand dairy industry. Therefore, caution should be exercised when generalising these findings to other herds or systems.

This study was also conducted over a relatively short time frame with a maximum of 2 seasons, which may have influence over genetic parameter estimation particularly as seen with PP heritability of 0.10. Longer-term datasets are generally preferred, as multi-year data improves stability of genetic parameter estimates and allows better capture of environmental variation across seasons (Haile-Mariam & Pryce, 2017). Longer-term datasets spanning numerous seasons would provide greater estimate stability and allow a more accurate assessment of environmental variation between years. Additionally, reliance on test-day records instead of whole-lactation measures (recording every milking), means some aspects of production performance may not be fully represented within the dataset recorded. While test-day models are widely used, full lactation records or more frequent sampling may provide a more complete representation of production performance over time (Sneddon et al., 2015a).

Another important limitation relates to the absence of direct BCS measurements within the dataset. While LWT remains a useful proxy for body size and weight fluctuations, it does not distinguish between structural mass and body fat reserves. This limits the ability to fully interpret energy balance dynamics and their relationship with production traits (Pryce & Harris, 2006; Roche et al., 2009). Inclusion of BCS records would significantly strengthen the biological interpretation of results and provide more in-depth insight into energy allocation across lactation.

Finally, even though pedigree-based genetic analyses were conducted, this study did not incorporate genomic information. Given the increasing importance of genomic selection in dairy breeding, the implementation of SNP or sequence data would have provided an immensely more detailed understanding of the genetic architecture underlying the observed traits (Hayes et al., 2010; Jiang et al., 2018). Future research integrating phenotypic, pedigree, and genomic data would allow more precise identification of the genetic relationships and improve the application of findings into breeding programmes.

Despite these limitations, the study provides a robust assessment of the relationship among LWT and milk production traits within a real-world pasture-based system. The consistency of findings across multiple analytical approaches supports the validity of the conclusions and provides a strong foundation for future research in this area.

5.10 Conclusion

This study provides insight into how LWT and milk production traits interact within a pasture-based dairy system, demonstrating that production outcomes emerge from coordinated biological processes rather than isolated traits. While larger cows exhibited greater overall production capacity, this advantage was not linear, with diminishing returns observed as body size increased. Highlighting the importance of taking into consideration not solely what an animal can produce, but how efficiently they can convert available resources into saleable output.

A key finding was the persistence of the negative relationship between MY and milk composition concentrations, even after accounting for major sources of variation such as age, lactations stage and live weight. Confirming the dilution of milk components is not simply a product of management or maturity but reflects fundamental physiological constraints underlying milk synthesis. As a result, strategies that focus solely on increasing milk volume risk reducing overall production efficiency when evaluated in terms of milksolids.

The use of repeated test-day records strengthened the ability to capture how production changes within individual cows over time, revealing that milk yield follows a structured and continuous trajectory across lactation. This reinforces the importance of treating production as temporally dependent and emphasises the value of modelling approaches that account for this continuity when evaluating biological relationships.

Differences across age groups and breeds further demonstrated that production is shaped by competing biological demands and genetic influences. Younger cows operate under constraints associated with growth, while older cows reach a plateau of stabilised performance once developmental demands decline. Breed comparisons illustrated contrasting production strategies, with some breeds favouring volume and others prioritising greater milksolids concentration with higher fat and protein percentages, with crossbreds offering a practical balance between the parental breeds. These distinctions indicate that no single production profile is universally optimal; instead, suitability depends on the production system and its constraints.

Together, these findings align with the objectives of pasture-based dairy systems where efficiency of milksolids production is prioritised over total volume. This underlines the value of selection approaches

that aim to balance body size, production capacity, and compositional output, rather than maximising individual traits in isolation, thus supporting both economic and biological sustainability.

This study additionally illustrates the potential of leveraging routinely collected data, such as LWT and milk records, to improve both breeding decisions and on-farm management. As precision equipment becomes more accessible, the ability to monitor individual animal performance over time will further enhance decision-making and allowing more targeted management strategies.

Although the observed relationships between LWT, milk production traits, and milk composition were based on a single herd and a limited timeframe, the consistency of patterns observed indicates the relationships identified are robust and reflect the underlying biological processes. Expanding future work to include longer-term and more densely sampled datasets, spanning multiple production seasons, lactations, and herds, together with genomic information, would strengthen understanding of these relationships significantly and support more precise selection strategies. In summary, this study demonstrates that effective dairy production relies on balancing biological capacity with efficiency within the constraints of the production system. Recognition and management of the trade-offs between milk yield, body size and condition, and milk composition are central to achieving sustainable and profitable outcomes in pasture-based dairy farming.

References

Abdalla, I. M, Hui, J, Nazar, M, Arbab, A. A. I, Xu, T, Abdu, S. M. N, Mao, Y, Yang, Z, & Lu, X. (2023). Identification of candidate genes and functional pathways associated with body size traits in Chinese Holstein cattle based on GWAS analysis. Animals, 13(6), 992. https://doi.org/10.3390/ani13060992

Alawneh, J. I., Stevenson, M. A., Williamson, N., Lopez-Villalobos, N., and Otley, T. (2012). The effect of liveweight change on reproductive performance in a seasonally calving, pasture fed dairy herd. Livestock Science 145 (1-3) 131-139. https://doi.org/10.1016/j.livsci.2011.12.025

Alemu, S. W., Handcock, R. C., & Garrick, D. J. (2025). Estimated genetic parameters for mature cow live weights during lactation in a multi-breed dairy cattle population. *New Zealand Journal of Agricultural Research*, 68(3), 547–554. https://doi.org/10.1080/00288233.2023.2261875

Alim, M. A., Dong, T., Xie, Y., Wu, X. P., Zhang, Y., Zhang, S., & Sun, D. X. (2014). Effect of polymorphisms in the CSN3 (κ-casein) gene on milk production traits in Chinese Holstein Cattle. *Molecular biology Reports*, 41(11), 7585–7593. https://doi.org/10.1007/s11033-014-3648-x

Ardicli, S, Samli, H, Vatansever, B, Soyudal, B, Dincel, D, & Balci, F. (2019). Comprehensive assessment of candidate genes associated with fattening performance in Holstein-Friesian bulls. *Archives Animal Breeding*, 62(1), 9–32. https://doi.org/10.5194/aab-62-9-2019

Auldist, M. J., O'Brien, G., Cole, D., Macmillan, K. L., & Grainger, C. (2007). Effects of varying lactation length on milk production capacity of cows in pasture-based dairying systems. *Journal of Dairy Science*, *90*(7), 3234–3241. https://doi.org/10.3168/jds.2006-683

Bauman, D. E., & Currie, W. B. (1980). Partitioning of nutrients during pregnancy and lactation: a review of mechanisms involving homeostasis and homeorhesis. *Journal of Dairy Science*, *63*(9), 1514–1529. https://doi.org/10.3168/jds.s0022-0302(80)83111-0

Berry, D. P, Lee, J. M, Macdonald, K. A, Stafford, K. J, Matthews, L. R, & Roche, J. R. (2007). Associations among body condition score, body weight, somatic cell count, and clinical mastitis in seasonally calving dairy cattle. *Journal of Dairy Science*, 90(2), 637–648. https://doi.org/10.3168/jds.s0022-0302(07)71546-1

Berry, D. P., & McCarthy, J. (2021). Contribution of genetic variability to phenotypic differences in on-farm efficiency metrics of dairy cows based on body weight and milksolids yield. *Journal of Dairy Science*. https://doi.org/10.3168/jds.2021-20542

Blott, S., Kim, J. J., Moisio, S., Schmidt-Küntzel, A., Cornet, A., Berzi, P., Cambisano, N., Ford, C., Grisart, B., Johnson, D., Karim, L., Simon, P., Snell, R., Spelman, R., Wong, J., Vilkki, J., Georges, M., Farnir, F., & Coppieters, W. (2003). Molecular dissection of a quantitative trait locus: a phenylalanine-to-tyrosine substitution in the transmembrane domain of the bovine growth hormone receptor is associated with a major effect on milk yield and composition. *Genetics*, 163(1), 253–266. https://doi.org/10.1093/genetics/163.1.253

Bourrat, P, & Lu, Q. (2017). Dissolving the missing heritability problem. *Philosophy of Science*, 84(5), 1055–1067. https://doi.org/10.1086/694007

Buckley, F., Dillon, P., Rath, M., & Veerkamp, R. F. (2000). The relationship between genetic merit for yield and live weight, condition score, and energy balance of spring calving Holstein Friesian dairy cows on grass based systems of milk production. *Journal of Dairy Science*, 83(8), 1878–1886. https://doi.org/10.3168/jds.S0022-0302(00)75060-0

Chen, L., Pryce, J. E., Hayes, B. J., & Daetwyler, H. D. (2021). Investigating the Effect of Imputed Structural Variants from Whole-Genome Sequence on Genome-Wide Association and Genomic Prediction in Dairy Cattle. Animals : an open access journal from MDPI, 11(2), 541. https://doi.org/10.3390/ani11020541

Cohen-Zinder, M, Seroussi, E, Larkin, D. M, Loor, J. J, Wind, A. E.-V. D, Lee, J.-H, Drackley, J. K, Band, M. R, Hernandez, A. G, Shani, M, Lewin, H. A, Weller, J. I, & Ron, M.. (2005). Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Research*, 15(7), 936–944. https://doi.org/10.1101/gr.3806705

DairyNZ. (2025). National dairy statistics. Retrieved January 29, 2026, from https://connect.dairynz.co.nz/2024-25_Dairy_Statistics/national-dairy-statistics.html

DairyNZ, & Livestock Improvement Corporation (LIC). (2025). New Zealand dairy statistics 2024/25. DairyNZ & LIC. https://www.dairynz.co.nz/resources/resource-list/new-zealand-dairy-statistics-2024-25-pdf/

Dechow, C. D, Rogers, G. W, & Clay, J. S. (2001). Heritabilities and correlations among body condition scores, production traits, and reproductive performance. *Journal of Dairy Science*, 84(1), 266–275. https://doi.org/10.3168/jds.s0022-0302(01)74476-1

Dechow, C. D, Rogers, G. W, Sander-Nielsen, U, Klei, L, Lawlor, T. J, Clay, J. S, Freeman, A. E, Abdel-Azim, G, Kuck, A, & Schnell, S. (2004). Correlations among body condition scores from various sources, dairy form, and cow health. *Journal of Dairy Science*, 87(10), 3526–3533. https://doi.org/10.3168/jds.s0022-0302(04)73489-x

Domecq, J. J, Skidmore, A. L, Lloyd, J. W, & Kaneene, J. B. (1997). Relationship Between Body Condition Scores and Milk Yield in a Large Dairy Herd of High Yielding Holstein Cows. *Journal of Dairy Science*, *80*(1), 101–112. https://doi.org/10.3168/jds.s0022-0302(97)75917-4

Fang, L, Sahana, G, Ma, P, et al., (2017). Exploring the genetic architecture and improving genomic prediction accuracy for mastitis and milk production traits in dairy cattle. Genetics Selection Evolution, 49, 44. https://doi.org/10.1186/s12711-017-0319-0

Fonterra Co-operative Group Limited. (2025). Farmgate milk price manual: FY26 (1 August 2025). https://www.fonterra.com

Friggens, N. C, Berg, P, Theilgaard, P, Korsgaard, I. R, Ingvartsen, K. L, Løvendahl, P, & Jensen, J. (2007). Breed and parity effects on energy balance profiles through lactation: evidence of

genetically driven body energy change. *Journal of Dairy Science*, *90*(11), 5291–5305. https://doi.org/10.3168/jds.2007-0173

Fulkerson, W. J, Davison, T. M, Garcia, S. C, Hough, G, Goddard, M. E, Dobos, R, & Blockey, M. (2008). Holstein-Friesian dairy cows under a predominantly grazing system: Interaction between genotype and environment. *Journal of Dairy Science,* *91*(2), 826–839. https://doi.org/10.3168/jds.2007-0147

Grisart, B., Coppieters, W., Farnir, F., Karim, L., Ford, C., Berzi, P., Cambisano, N., Mni, M., Reid, S., Simon, P., Spelman, R., Georges, M., & Snell, R. (2002). Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Research*, 12(2), 222–231. https://doi.org/10.1101/gr.224202

Handcock, R. C, Lopez-Villalobos, N, McNaughton, L. R, Back, P. J, Edwards, G. R, & Hickson, R. E. (2019). Live weight and growth of Holstein-Friesian, Jersey and crossbred dairy heifers in New Zealand. *New Zealand Journal of Agricultural Research*, *62*(2), 173–183. https://doi.org/10.1080/00288233.2018.1465984

Haile-Mariam, M., & Pryce, J. E. (2017). Genetic parameters for lactose and its correlation with other milk production traits and fitness traits in pasture-based production systems. *Journal of Dairy Science*, 100(5), 3754–3766. https://doi.org/10.3168/jds.2016-11952

Hardie, L. C., VandeHaar, M. J., Tempelman, R. J., Weigel, K. A., Armentano, L. E., Wiggans, G. R., Veerkamp, R. F., de Haas, Y., Coffey, M. P., Connor, E. E., Hanigan, M. D., Staples, C., Wang, Z., Dekkers, J. C. M., & Spurlock, D. M. (2017). The genetic and biological basis of feed efficiency in mid-lactation Holstein dairy cows. *Journal of Dairy Science*, 100(11), 9061–9075. https://doi.org/10.3168/jds.2017-12604

Hayes, B. J, & Daetwyler, H. D. (2019). 1000 bull genomes project to map simple and complex genetic traits in cattle: Applications and outcomes. *Annual Review of Animal Biosciences*, *7*(1), 89–102. https://doi.org/10.1146/annurev-animal-020518-115024

Hayes, B. J, Pryce, J, Chamberlain, A. J, Bowman, P. J, & Goddard, M. E. (2010). Genetic architecture of complex traits and accuracy of genomic prediction: Coat colour, milk-fat percentage, and type in holstein cattle as contrasting model traits. *PLOS Genetics*, *6*(9), e1001139. https://doi.org/10.1371/journal.pgen.1001139

Hellwege, J. N, Keaton, J. M, Giri, A, Gao, X, Velez Edwards, D. R, & Edwards, T. L.. (2017). Population Stratification in Genetic Association Studies. *Current Protocols in Human Genetics*, *95*(1), 1.22.1–1.22.23. https://doi.org/10.1002/cphg.48

Horan, B, Dillon, P, Faverdin, P, Delaby, L, Buckley, F, & Rath, M. (2005). The interaction of strain of Holstein-Friesian cows and pasture-based feed systems on milk yield, body weight, and body condition score. *Journal of Dairy Science*, *88*(3), 1231–1243. https://doi.org/10.3168/jds.S0022-0302(05)72790-9

Jagusiak, W., Ptak, E., Otwinowska-Mindur, A., & Zarnecki, A. (2023). Genetic relationships of body condition score and locomotion with production, type and fertility traits in Holstein-Friesian cows. Animal : an international journal of animal bioscience, 17(6), 100816. https://doi.org/10.1016/j.animal.2023.100816

Jayawardana, J. M. D. R., Lopez-Villalobos, N., McNaughton, L. R., & Hickson, R. E. (2023). Heritabilities and genetic and phenotypic correlations for milk production and fertility traits of spring-calved once-daily or twice-daily milking cows in New Zealand. *Journal of Dairy Science*, 106(3), 1910–1924. https://doi.org/10.3168/jds.2022-22431

Jiang, J., Cole, J. B., Da, Y., VanRaden, P. M., & Ma, L. (2018). Fast Bayesian fine-mapping of 35 production, reproduction and body conformation traits in dairy cattle. bioRxiv, 428227. https://doi.org/10.1101/428227

Kamprasert, N., Aliloo, H., van der Werf, J. H. J., & Clark, S. A. (2024). Short communication: Accuracy of whole-genome sequence imputation in Angus cattle using within-breed and multi breed reference populations. *Animal*, *18*(3), 101087. https://doi.org/10.1016/j.animal.2024.101087

Krattenmacher, N., Thaller, G., & Tetens, J. (2019). Analysis of the genetic architecture of energy balance and its major determinants dry matter intake and energy-corrected milk yield in primiparous Holstein cows. *Journal of Dairy Science*, 102(4), 3241–3253. https://doi.org/10.3168/jds.2018-15480

Ledinek, M., Gruber, L., Steininger, F., Zottl, K., Royer, M., Krimberger, K., … Fuerst-Waltl, B. (2019). Analysis of lactating cows in commercial Austrian dairy farms: diet composition, and influence of genotype, parity and stage of lactation on nutrient intake, body weight and body condition score. Italian Journal of Animal Science, 18(1), 202–214. https://doi.org/10.1080/1828051X.2018.1504632

Lembeye, F., Lopez-Villalobos, N., Burke, J. L., & Davis, S. R. (2016). Estimation of genetic parameters for milk traits in cows milked once- or twice-daily in New Zealand. Livestock Science, 185, 142–147. https://doi.org/10.1016/j.livsci.2016.01.022

Li, B., Fikse, W. F., Løvendahl, P., Lassen, J., Lidauer, M. H., Mäntysaari, P., & Berglund, B.. (2018). Genetic heterogeneity of feed intake, energy-corrected milk, and body weight across lactation in primiparous Holstein, Nordic Red, and Jersey cows. *Journal of Dairy Science*, *101*(11), 10011–10021. https://doi.org/10.3168/jds.2018-14611

Littlejohn, M. D, Tiplady, K, Lopdell, T, Law, T. A, Scott, A, Harland, C, Sherlock, R, Henty, K, Obolonkin, V, Lehnert, K, Macgibbon, A, Spelman, R. J, Davis, S. R, & Snell, R. G. (2014). Expression variants of the lipogenic AGPAT6 gene affect diverse milk composition phenotypes in Bos taurus. *PLoS ONE*, *9*(1), e85757. https://doi.org/10.1371/journal.pone.0085757

Lopdell, T. J, Tiplady, K, Struchalin, M, Johnson, T. J. J, Keehan, M, Sherlock, R, Couldrey, C, Davis, S. R, Snell, R. G, Spelman, R. J, & Littlejohn, M. D. (2017). DNA and RNA-sequence based GWAS

highlights membrane-transport genes as key modulators of milk lactose content. *BMC Genomics*, *18*(1), 968. https://doi.org/10.1186/s12864-017-4320-3

Lopez-Villalobos, N., Spelman, R. J., Melis, J., Davis, S. R., Berry, S. D., Lehnert, K., Sneddon, N. W., Holroyd, S. E., MacGibbon, A. K., & Snell, R. G. (2020). Genetic correlations of milk fatty acid contents predicted from milk mid-infrared spectra in New Zealand dairy cattle. *Journal of Dairy Science*, *103*(8), 7238–7248. https://doi.org/10.3168/jds.2019-17971

Livestock Improvement Corporation (LIC). (2025). Dairy statistics. Retrieved November 10, 2025, from https://www.lic.co.nz/about/research-publications/dairy-statistics/

McClearn, B., Delaby, L., Gilliland, T. J., Guy, C., Dineen, M., Coughlan, F., Buckley, F., & McCarthy, B. (2020). An assessment of the production, reproduction, and functional traits of Holstein-Friesian, Jersey × Holstein-Friesian, and Norwegian Red × (Jersey × Holstein-Friesian) cows in pasture-based systems. *Journal of Dairy Science*, 103(6), 5200–5214. https://doi.org/10.3168/jds.2019-17476

McCarthy, S., Berry, D. P., Dillon, P., Rath, M., & Horan, B. (2007). Influence of Holstein-Friesian strain and feed system on body weight and body condition score lactation profiles. *Journal of Dairy Science*, 90(4), 1859–1869. https://doi.org/10.3168/jds.2006-501

Maiorano, A. M., Lourenco, D. L., Tsuruta, S., Ospina, A. M. T., Stafuzza, N. B., Masuda, Y., Filho, A. E. V., Cyrillo, J. N. D. S. G., Curi, R. A., & Silva, J. A. I. V. (2018). Assessing genetic architecture and signatures of selection of dual purpose Gir cattle populations using genomic information. *PLoS ONE*, 13(8), e0200694. https://doi.org/10.1371/journal.pone.0200694

Ma, Y., Khan, M. Z., Xiao, J., Alugongo, G. M., Chen, X., Chen, T., Liu, S., He, Z., Wang, J., Shah, M. K., & Cao, Z. (2021). Genetic Markers Associated with Milk Production Traits in Dairy Cattle. *Agriculture*, *11*(10), 1018. https://doi.org/10.3390/agriculture11101018

Miglior, F., Sewalem, A., Jamrozik, J., Bohmanova, J., Lefebvre, D. M., & Moore, R. K. (2007). Genetic analysis of milk urea nitrogen and lactose and their relationships with other production traits in canadian holstein cattle. *Journal of Dairy Science*, *90*(5), 2468–2479. https://doi.org/10.3168/jds.2006-487

Mullen, M. P., Berry, D. P., Howard, D. J., Diskin, M. G., Lynch, C. O., Giblin, L., Kenny, D. A., Magee, D. A., Meade, K. G., & Waters, S. M. (2011). Single Nucleotide Polymorphisms in the Insulin-Like Growth Factor 1 (IGF-1) Gene are Associated with Performance in Holstein-Friesian Dairy Cattle. *Frontiers in Genetics*, 2, 3. https://doi.org/10.3389/fgene.2011.00003

Neville M. C. (2001). Anatomy and physiology of lactation. Pediatric clinics of North America, 48(1), 13–34. https://doi.org/10.1016/s0031-3955(05)70283-2

Patton, J, Kenny, D. A, McNamara, S, Mee, J. F, O'Mara, F. P, Diskin, M. G, & Murphy, J. J. (2007). Relationships among milk production, energy balance, plasma analytes, and reproduction in Holstein-Friesian cows. *Journal of Dairy Science*, *90*(2), 649–658. https://doi.org/10.3168/jds.S0022-0302(07)71547-3

Pegolo, S., Yu, H., Morota, G., Bisutti, V., Rosa, G. J. M., Bittante, G., & Cecchinato, A. (2021). Structural equation modeling for unraveling the multivariate genomic architecture of milk proteins in dairy cattle. *Journal of Dairy Science*, *104*(5), 5705–5718. https://doi.org/10.3168/jds.2020-18321

Pereira Ciochetti, N, Lugli-Moraes, B, Santos Da Silva, B, & Rovaris, D. L.. (2023). Genome-wide association studies: utility and limitations for research in physiology. *The Journal of Physiology*, *601*(14), 2771–2799. https://doi.org/10.1113/jp284241

Prendiville, R, Pierce, K. M, & Buckley, F.. (2009). An evaluation of production efficiencies among lactating Holstein-Friesian, Jersey, and Jersey × Holstein-Friesian cows at pasture. *Journal of Dairy Science*, *92*(12), 6176–6185. https://doi.org/10.3168/jds.2009-2292

Price, A. L., Zaitlen, N. A., Reich, D., & Patterson, N. (2010). New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics*, *11*(7), 459–463. https://doi.org/10.1038/nrg2813

Pires, J. A., Delavaud, C., Faulconnier, Y., Pomiès, D., & Chilliard, Y. (2013). Effects of body condition score at calving on indicators of fat and protein mobilization of periparturient Holstein-Friesian cows. *Journal of Dairy Science*, 96(10), 6423–6439. https://doi.org/10.3168/jds.2013-6801

Pryce, J. E, & Harris, B. (2006). Genetics of Body Condition Score in New Zealand Dairy Cows. *Journal of Dairy Science*, *89*(11), 4424–4432. https://doi.org/10.3168/jds.s0022-0302(06)72490-0

Roche, J. R., Macdonald, K. A., Burke, C. R., Lee, J. M., & Berry, D. P. (2007). Associations among body condition score, body weight, and reproductive performance in seasonal-calving dairy cattle. *Journal of Dairy Science*, *90*(1), 376–391. https://doi.org/10.3168/jds.S0022-0302(07)72639-5

Roche, J. R., Friggens, N. C., Kay, J. K., Fisher, M. W., Stafford, K. J., & Berry, D. P. (2009). Invited review: Body condition score and its association with dairy cow productivity, health, and welfare. *Journal of Dairy Science*, 92(12), 5769–5801. https://doi.org/10.3168/jds.2009-2431

Schmidtmann, C., Segelke, D., Bennewitz, J., Tetens, J., & Thaller, G. (2023). Genetic analysis of production traits and body size measurements and their relationships with metabolic diseases in German Holstein cattle. *Journal of Dairy Science*, 106(1), 421–438. https://doi.org/10.3168/jds.2022-22363

Schopen, G. C, Visker, M. H, Koks, P. D, Mullaart, E, van Arendonk, J. A, & Bovenhuis, H. (2011). Whole-genome association study for milk protein composition in dairy cattle. *Journal of Dairy Science*, 94(6), 3148–3158. https://doi.org/10.3168/jds.2010-4030

Sieber, M, Freeman, A. E, & Kelley, D. H. (1988). Relationships between Body Measurements, Body Weight, and Productivity in Holstein Dairy Cows. *Journal of Dairy Science*, *71*(12), 3437–3445. https://doi.org/10.3168/jds.s0022-0302(88)79949-x

Si, Q. N. (2024). Genome-wide association studies for milk production in dairy cattle. Animal Molecular Breeding, 14(5), 130–140. https://doi.org/10.5376/amb.2024.14.0032

Sneddon, N., Lopez-Villalobos, N., Davis, S., Hickson, R. and Shalloo, L. (2015a), Genetic parameters for milk components including lactose from test day records in the New Zealand dairy herd. *New Zealand Journal of Agricultural Research*, 58: 97-107. https://doi.org/10.1080/00288233.2014.978482

Sneddon, N. W, Lopez-Villalobos, N, Hickson, R. E, Shalloo, L, Davis, S, Garrick, D. J, & Geary, U. (2015b). Predicted dairy product yields and deficits of lactose for manufacturing under differing selection and manufacturing scenarios in New Zealand. *New Zealand Journal of Agricultural Research*, 58(4), 432–440. https://doi.org/10.1080/00288233.2015.1073160

Sneddon, N., Lopez-Villalobos, N., Davis, S., Hickson, R., Shalloo, L., & Garrick, D. (2016). Estimates of genetic and crossbreeding parameters for milk components and potential yield of dairy products from New Zealand dairy cattle. *New Zealand Journal of Agricultural Research*, *59*(1), 79–89. https://doi.org/10.1080/00288233.2015.1131723

Sousa Junior, L. P. B., Pinto, L. F. B., Cruz, V. A. R., Oliveira Junior, G. A., Oliveira, H. R., Chud, T. S., Pedrosa, V. B., Miglior, F., Schenkel, F. S., & Brito, L. F. (2024). Genome-wide association and functional genomic analyses for body conformation traits in North American Holstein cattle. *Frontiers in Genetics*, 15, 1478788. https://doi.org/10.3389/fgene.2024.1478788

Spelman, R. J., & Garrick, D. J. (1997). Effect of live weight and differing economic values on responses to selection for milk fat, protein, volume, and live weight. *Journal of Dairy Science*, 80(10), 2557–2562. https://doi.org/10.3168/jds.S0022-0302(97)76211-8

Sun, R., Jiang, X., Hao, Y., Li, Y., Bai, Y., Xia, C., & Song, Y. (2025). Effect of body condition score loss during the transition period on metabolism, milk yield and health in Holstein cows. *Journal of Veterinary Research*, 69(1), 91–99. https://doi.org/10.2478/jvetres-2025-0004

Tam, V, Patel, N, Turcotte, M, Bossé, Y, Paré, G, & Meyre, D.. (2019). Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics*, *20*(8), 467–484. https://doi.org/10.1038/s41576-019-0127-1

Vance, E. R., Ferris, C. P., Elliott, C. T., Hartley, H. M., & Kilpatrick, D. J. (2013). Comparison of the performance of Holstein-Friesian and Jersey×Holstein-Friesian crossbred dairy cows within three contrasting grassland-based systems of milk production. *Livestock Science*, 151(1), 66-79. https://doi.org/10.1016/j.livsci.2012.10.011

Walsh, S., Buckley, F., Pierce, K., Byrne, N., Patton, J., & Dillon, P. (2008). Effects of breed and feeding system on milk production, body weight, body condition score, reproductive performance, and postpartum ovarian function. *Journal of Dairy Science*, 91(11), 4401–4413. https://doi.org/10.3168/jds.2007-0818

Wu, M., Lin, Z., Ma, S., Chen, T., Jiang, R., & Wong, W. H.. (2017). Simultaneous inference of phenotype-associated genes and relevant tissues from GWAS data via Bayesian integration of multiple tissue-specific gene networks. *Journal of Molecular Cell Biology*, *9*(6), 436–452. https://doi.org/10.1093/jmcb/mjx059

Yin, T., & König, S. (2019). Genome-wide associations and detection of potential candidate genes for direct genetic and maternal genetic effects influencing dairy cattle body weight at different ages. *Genetics Selection Evolution*, *51*(1), 4–4. https://doi.org/10.1186/s12711-018-0444-4

Zhang, H., Mi, S., Brito, L. F., Hu, L., Wang, L., Ma, L., Xu, Q., Guo, G., Yu, Y., & Wang, Y. (2023). Genomic and transcriptomic analyses enable the identification of important genes associated with subcutaneous fat deposition in Holstein cows. *Journal of Genetics and Genomics*, 50(6), 385–397. https://doi.org/10.1016/j.jgg.2023.01.011

Zhou, C., Li, C., Cai, W., Liu, S., Yin, H., Shi, S., Zhang, Q., & Zhang, S. (2019). Genome-Wide Association Study for Milk Protein Composition Traits in a Chinese Holstein Population Using a Single-Step Approach. *Frontiers in Genetics*, 10, 72. https://doi.org/10.3389/fgene.2019.00072