Improving the measurement of live weight and body condition score in sheep

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

Doctor of Philosophy

in

Animal Science

Massey University, Turitea, Palmerston North

New Zealand.

Jimmy Semakula

2021

Thesis Supervisors:

Chief supervisor:

Professor Paul Richard Kenyon

Co-supervisors:

Professor Hugh Thomas Blair

Dr Rene Anne Corner-Thomas

Professor Stephen Todd Morris

Abstract

Liveweight (LW) and body condition score (BCS) are important performance indicators in sheep management, providing a basis for decision making. Therefore, accurate measurement of these traits is imperative. The overall aims of this thesis were to: 1) Determine the factors affecting the rate of LW loss of fasting sheep, 2) derive equations to predict LW and LW change over a short time period (1 to 8 hours), 3) evaluate the factors affecting the relationship between ewe LW and BCS, and 4) derive equations for predicting ewe BCS. In the LW studies, lambs were offered three herbage availability levels (Low, Medium and High) in autumn or winter. Similarly, mixed-aged ewes at different physiological states were offered two herbage levels (Low or High). These studies were conducted in two stages: A) calibration stage and B) validation stage.

Equations to predict without delay LW were developed at the calibration stage and validated on data collected from independent ewes from different farms. The rate of ewe LW loss was influenced by herbage type and availability, and season. Further, in pregnant ewes, liveweight loss was influenced by stage of pregnancy, but not pregnancy-rank. Applying correction equations improved the prediction accuracy of without delay LW estimates up to 55% and 69% in ewe lambs and mixed aged ewes compared with using the delayed weights, respectively.

For the BCS studies, LW and BCS data of ewes were collected at regular times of the annual production cycle until they were six years of age. Using a ewe's LW and BCS records to predict their current BCS using a linear model gave moderately accurate estimates. A different dataset, which included foetal- and fleece weight-adjusted LW and height at withers was then used. It was found that equations combining LW, LW change and previous BCS explained more variability in current BCS and were more accurate than LW-alone based models but the addition of adjusted LW and height at withers gave no further benefit to the BCS prediction models. Applying machine learning classification algorithms such as extreme gradient boosted trees and Random forest on a 3-point BCS scale achieved very good BCS prediction accuracies (> 85%).

These combined findings provide useful prediction equations that could be incorporated into weighing systems, which along with EID would improve sheep production by aiding management decision making.

Acknowledgements

I can do all things through Christ who strengthens me (Philippians 4:13). I thank the Almighty God who has granted me the abilities and skills to make things possible.

In a special way, my appreciation goes to my supervisor Prof. Paul R Kenyon. You have been so amazing to me. Thank you for mentoring me into an independent researcher. Your motivation, encouragement and counsel on this research has been priceless. I would also like to appreciate my co-supervisors Prof Hugh T Blair, Prof Stephen T Morris and Dr Rene A Corner-Thomas for their continuous support, guidance, and immense knowledge. You have been a great team, people of talent; may you keep up the spirit.

My sincere thanks also go to the staff of the department of Animal Science: Mr Dean Burnham, Dr Catriona Jenkinson, and Geoff Purchas. I appreciate your dedication and assistance in various ways during the different phases of my study. I would like to thank Mr Dean Burnham in a special way. I spent most of my time with you in the field doing banal data collecting and you shared your sheep husbandry knowledge without hesitation. Those memories will always endure.

A special thanks to my family, especially my parents Samuel Semakula and Grace Akwi Semakula, my siblings (Cate, Ronald, and Moses). Your prayers were what kept me going through this study. Lastly, but more importantly, my deepest appreciation to Hanifah Nagayi “our mom” and my children (Drusilla and Moses) for your love, patience and encouragement (“for daddy to complete his class- and homework”) through it all. I love you and thank God for you.

Table of contents

Table of contents

List of tables

List of figures

Abbreviations

ADF:	Acid detergent fibre
Adj.R^2:	Adjusted coefficient of determination
AIC:	Akaike's information criterion
ANN:	Artificial neural networks
BCS:	Body condition score
CART:	Classification and regression trees
R^2:	Coefficient of determination (multiple regression)
r^2:	Coefficient of determination (simple regression)
CCC:	Lin's concordance correlation coefficient
CV:	Coefficient of variation
CP:	Crude protein
DM:	Dry matter
XGB:	Extreme gradient boosted trees
FN:	False negatives
FP:	False positives
LM:	General linear model
GLS:	Generalised least squares
HW:	Height at withers
KNN:	K-nearest neighbours
LDA:	Linear discriminant analysis
LMM:	Linear mixed-effects model
LW:	Liveweight
ML:	Machine learning
MUAEC:	Massey University Animal Ethics committee
ML:	Maximum likelihood
MAE:	Mean absolute error
MAPE:	Mean absolute percent error
ME:	Metabolizable energy
MCA:	Multiple correspondence analysis
NIRS:	Near-infrared reflectance spectroscopy

NDF:	Neutral detergent fibre
OMD:	Organic matter digestibility
P100:	100 days of pregnancy
P130:	130 days of pregnancy
PCA:	Principal component analysis
PR:	Pregnancy-rank
QUAD:	Quadratic
RFID:	Radio frequency identification
RF:	Random forests
RPE:	Relative percent error
RPD:	Residual prediction deviation
RPIQ:	Ratio of performance to interquartile distance
ReML:	Restricted maximum likelihood
RMSE:	Root mean square error
SQRT:	Square root
SVM:	Support vector machines
TP:	True positives
WSC:	Water soluble carbohydrate

Chapter 1. General Introduction

The New Zealand sheep industry relies on extensive grazing pastoral systems with average flock sizes greater than 2500 sheep (Cranston et al., 2017). Ewes undergo an annual production cycle with four critical stages of economic importance at which critical management decisions are made. Key management decisions are made about nutrition, reproduction and health for improved ewe performance and productivity. Such management decisions should be based on credible and accurate data. For example, inaccurate liveweights could lead to poor decisions when a comparison of liveweights is required.

Liveweight (LW) is a broadly accepted proxy for the energy status of sheep at a given time, while change in liveweight is indicative of whether it is in either a positive energy balance (liveweight gain) or a negative energy balance (liveweight loss) (Young and Corbett, 1972; Brown et al., 2005; Wishart et al., 2017). Ewe management decision making is based on performance target thresholds and optimal ranges around these targets. For example, the threshold breeding liveweight for ewe lambs to reach puberty is between 40% and 70% of their mature liveweight (Dyrmundsson, 1973; Jainudeen et al., 2000). Further, several studies have reported a positive relationship between ewe reproductive rate with liveweight which becomes less significant after reaching an optimum threshold weight (Rutherford et al., 2003; Kenyon et al., 2004b; Corner-Thomas et al., 2015b). Therefore, it is imperative that accurate liveweight data measurement is achieved.

Ewe liveweight is relatively stable over a short period of time (a few minutes), but alters over longer time periods in response to environmental and physiological conditions (Coates and Penning, 2000b; Wishart et al., 2017). The accuracy of liveweight measurements is affected by a number of factors including: gut-fill (digesta and urine), growth, nutrition, health, stress, fleece weight, physiological state and genotype (Kenyon et al., 2014; Brown et al., 2015). The contents of the rumen (fluid and feed) can account for between 10 and 23% of total liveweight in ruminants (Hughes, 1976; Moyo and Nsahlai, 2018). Liveweight fluctuations due to gut-fill in ruminants can be affected by factors influencing feed intake such as age and size of the animal, time of day, ambient temperature, grazing behaviour and time since last meal (Hughes, 1976; Coates and Penning, 2000b; Hogan et al., 2007; Burnham et al., 2009; Gregorini, 2012; Wilson et al., 2015; Wishart et al., 2017).

Automatic weighing systems can record up to 400 weights per hour without interruptions (https://www.livestock.tru-test.com), thus, requiring six to seven hours to weigh an average New Zealand flock (2500 sheep). Further, mustering and routine on-farm sheep handling in addition to weighing can increase the length of time sheep are restricted from accessing feed and water supplies while waiting to be weighed. Delays in weighing ewes can lead to weight loss due to a reduction in gut-fill and body fluids (Hogan et al., 2007; Burnham et al., 2009; Wilson et al., 2015). Varying levels of weight loss have been reported within flocks waiting to be weighed. In ewe lambs, losses between 1.8% and 9.8% of initial liveweight after six hours have been reported (Hughes, 1976; Burnham et al., 2009; Wishart et al., 2017). In mature ewes, losses of 1.78 kg (2.7% of initial liveweight) and 1.69 kg (2.6% of initial liveweight) in single and twin bearing ewes at day130 of pregnancy after six hours, and 3.4 kg (5.3% of initial liveweight) and 2.9 kg (4.5% of initial liveweight) after 12 hours have been reported (Burnham et al., 2009). These levels of liveweight loss can significantly interfere with the accuracy of comparison of liveweights, and changes in liveweight over time.

Pre-fasting gut-fill is important in determining the rate of sheep liveweight loss during fasting (Kirton et al., 1968; Kirton et al., 1971; Thompson et al., 1987). The degree of gut-fill, retention time of particles in the gastrointestinal tract and passage rate can be affected by the quality and quantity (including particle size and consistence) of dry matter intake in ruminants (Alwash and Thomas, 1971; Haaland and Tyrrell, 1982; Varga and Prigge, 1982; Kaske and Groth, 1997). Therefore, it is likely that differences in herbage type and availability offered to sheep can result in variation in the rate of liveweight loss during fasting.

Several strategies can be used to reduce variability in liveweight including removal of feed and water for fixed periods of time prior to weighing, standardizing weighing procedures, taking multiple liveweights readings per individual per day over successive days, weighing at a specific time relative to sunrise, standardizing the feed offered prior to weighing and/or increasing the number of animals and repetitions of a study (Coates and Penning, 2000b; Wishart et al., 2017). Implementing such methodologies to reduce variation, however, can be time-consuming and therefore, are not generally utilized, except in experimental situations. Thus, there is a need for a new approach to determine and adjust for variations in liveweight among animals, across time. The on-going improvements in weighing equipment, software and data management (Brown et al., 2015; Wilson et al., 2015) may offer a solution, as there is capacity for the time stamping of individual animal weights. To date, technology companies have not yet incorporated weighing methodology in their systems to deal with this variation. Using the time at which animals were collected for weighing in equations for predicting liveweight change post removal from feed, makes it possible to calculate more consistent liveweight measurements (Burnham et al., 2009; Wilson, 2014; Wilson et al., 2015; Wishart et al., 2017). Modern weighing systems should be able to provide this information to farmers instantaneously. However, liveweight loss and thus, the methodologies to adjust for liveweight are likely to differ according to herbage type and availability, season and ewe physiological state and their interactions, but it is not clear to what extent. Therefore, there is need to interrogate if these factors affect the rate of liveweight loss of ewes and, if they do, methodologies need to be developed to adjust liveweight for these factors.

Body condition score is a subjective measure which provides an estimate of an animal's soft tissue reserves, predominantly fat, that can be used by farmers and researchers to determine the physiological state of an animal (Morris et al., 2002; Vieira et al., 2015). Like LW, BCS is related to ewe production and reproductive traits and there are thresholds or ranges of BCS values that are optimal for productivity (Kenyon et al., 2004a; Kleemann and Walker, 2005; Scaramuzzi et al., 2006; Kenyon et al., 2011b; Kenyon et al., 2014). Body condition score can circumvent the shortcomings of liveweight (LW) mentioned above. Further, body condition score can be easily learned and is cost-effective and requires no specialized equipment (Kenyon et al., 2014). Despite the advantages of using BCS over liveweight to better manage flocks, it is uncommon (7−40%) for farmers, especially in extensive production systems, to regularly and objectively do so (Jones et al., 2011). The reasons for low BCS uptake among farmers include the subjective nature, labour burden and constant recalibration of assessors (Kenyon et al., 2014). Strategies to increase the adoption and use of BCS among farmers and the reliability of measures have been limited to promotional workshops and hands-on training (Kenyon et al., 2014). However, these strategies do not directly address how to reduce the labour burden associated with hands-on BCS. Therefore, it is argued that, consistent and accurate alternative methods to estimate body condition score of sheep that require less hands-on measurement would likely be advantageous and improve uptake and use. To date there are no known attempts to exploit the relationship between liveweight and BCS to predict the later. This thesis aims to unlock the potential of exploiting the relationship between LW and BCS to allow both measurements automatically recorded on a single weighing head screen. The aims of this thesis, therefore, are firstly to determine the factors affecting the rate of LW loss of fasting ewes, 2) derive equations to predict LW and LW change over the short term (1 to 8 hours), 3) evaluate the factors affecting the relationship between ewe LW and BCS, 4) derive equations predicting ewe current ewe BCS.

Specific objectives of this thesis were to:

1) Determine the effect of feed type on the rate of ewe lamb liveweight loss during fasting (Chapter 3)

2) Determine the effect of herbage availability and season on the rate of ewe lamb liveweight loss during fasting (Chapter 4)

3) Determine the effect of herbage availability, physiological state (non-pregnant or pregnant), stage of pregnancy (100 or 130 days of pregnancy) and pregnancy-rank (single- or twin-bearing) on the rate of mixed-aged ewe liveweight loss during fasting (Chapter 5)

4) Determine the effect of ewe age, stage of annual production cycle and pregnancy-rank on the relationship between liveweight and body condition score (Chapter 6)

5) Predict the current body condition score from a ewe's liveweight, liveweight change and previous body condition record (Chapter 7)

6) Determine if machine learning algorithms could be a better alternative to the linear model in predicting ewe BCS from liveweight records (Chapter 8)

7) Determine if using adjusted liveweight, liveweight change, previous BCS and height at withers would improve the accuracy of current ewe BCS prediction (Chapter 9)

Chapter 2. Literature review

2.1 Preamble

Liveweight (LW) and body condition score (BCS) are indicators of the body condition and body reserves providing a basis for management decisions of sheep. Therefore, it is imperative that they are measured accurately. Liveweight affects productivity and these relationships are summarised in this review. Accurate measurement of liveweight depends on the instrument used, animal factors, environmental factors and human factors (Wilson, 2014; Elwood, 2017). With the advent of automatic electronic weighing systems, potential error due to human effects and instrumentation is becoming obsolete leaving animal and environmental factors as having the greatest effect on the accuracy of a given weight measurement (Wilson, 2014; Wilson et al., 2015). Accordingly, this literature review focuses on the animal factors (predominantly in sheep) affecting liveweight measurement with an emphasis on gut-fill (fluid and feed). Gut-fill can account for between 10 and 23% of total liveweight in ruminants (Hughes, 1976; Moyo and Nsahlai, 2018). Factors affecting gut-fill will thus be reviewed. Body condition score measurement is considered more reliable than liveweight as it circumvents the factors that compromise the accuracy of liveweight measurement (van Burgel et al., 2011; Kenyon et al., 2014; Brown et al., 2015; Morel et al., 2016). However, BCS measurement is a subjective (hands-on) method for assessing animal performance (Russel et al., 1969; Morris et al., 2002; van Burgel et al., 2011; Kenyon et al., 2014). Reliability and repeatability of BCS measurements are of concern (Evans, 1978; Calavas et al., 1998; Curnow et al., 2011; van Burgel et al., 2011; Phythian et al., 2012; Kenyon et al., 2014). Therefore, this review focuses on factors affecting sheep BCS measurement reliability and repeatability. Further, BCS is not popular among most farmers because it can be labour intensive and requires training (Jones et al., 2011; Corner-Thomas et al., 2016). For this reason, a review has been done on probable indicators (proxy variables) of BCS and research has been undertaken to study the possibility of indirectly predicting BCS. Lastly, the importance of liveweight and BCS in sheep productivity, the various methods of measuring liveweight and BCS, advantages, and disadvantages associated with them will also be discussed. This review concentrates specifically on sheep, however, where appropriate, references have been made to other species.

2.2 Effect of liveweight in sheep productivity

The effects of liveweight on the performance of a ewe and its progeny are well documented (Ferguson et al., 2011; Oldham et al., 2011; Hickson et al., 2012; Kenyon et al., 2012a; Brown et al., 2015) and will be briefly reviewed. The review section is a summary of these relationships.

2.2.1 Breeding, fertility, productive rates

Liveweight of ewes affects attainment of puberty, fertility (pregnant ewes per 100 ewes exposed to rams) and productive rates (foetuses *in utero* per 100 ewes exposed to rams) of ewes, which impacts its productivity (Newton et al., 1980; Stephenson et al., 1980; Smith, 1982; Saul et al., 2011; Brown et al., 2015). Previous studies have demonstrated the relationship between liveweight and the reproductive traits of puberty onset (Meyer and French, 1979; McMillan and Moore, 1983; Rosales et al., 2013), ovulation rate (Michels et al., 2000; Kenyon et al., 2004b; Kleemann and Walker, 2005; Scaramuzzi et al., 2006; Rosales et al., 2013), conception and multiple birth rate (Kenyon et al., 2004b; Kenyon et al., 2014; Aktaş et al., 2015) and lamb growth and survival (Hinch et al., 1985; Oldham et al., 2011; Kenyon et al., 2014; Aktaş et al., 2015). A summary of the studies on the relationship between liveweight and reproductive performance of ewes from puberty onset to pregnancy is given in Table 2.1.

Liveweight is a major factor influencing puberty onset in sheep (Quirke et al., 1985; Khalifa et al., 2013; Zarkawi and Al-Daker, 2016). A positive relationship between liveweight and time to puberty has been demonstrated (Ferra et al., 2010), with the threshold breeding liveweight for ewe lambs to reach puberty being between 40% and 70% of their mature liveweight (Hafez, 1952; Dyrmundsson, 1973; Jainudeen et al., 2000). Furthermore, the heavier ones within a flock are more likely to show oestrus and successfully join as lambs compared with their lighter contemporaries (Kenyon et al., 2010). Heavier ewe lambs and adult ewes are more likely to mate in the first 17 days of the breeding period and are more likely to have multiple offspring (Kenyon et al., 2004a; Kenyon et al., 2005; Kenyon et al., 2006). Liveweight at mating has also been reported to have positive effects on the proportion of ewe lambs displaying oestrus (Meyer and French, 1979; McMillan and Moore, 1983; Kenyon et al., 2005; Kenyon et al., 2006).

Ewe ovulation rate is a major driver of ewe fecundity and is sensitive to liveweight (Rhind et al., 1984a; Rowe, 2003). Liveweight during breeding has been reported to be positively associated with increased ovulation rates in both ewe lamb and adult ewes (Morley et al., 1978; Kenyon et al., 2004b; Scaramuzzi et al., 2006). Morley et al. (1978) working on a wide range of sheep genotypes reported an average increase of 2% in ovulation rates for every 1 kg increase in liveweight while Edey (1968), working on Peppin Merinos between 35 and 53.5 kg, reported a 2-5% increase per 2.5 kg liveweight change. Kenyon et al. (2004b) reported that ovulation rates increased with liveweight, plateauing after 62.6 kg and 48.5 kg in mixed aged Romney and two-tooth composite ewes respectively. Rutherford et al. (2003), reported that in mature predominantly Coopworth ewes, any increase in liveweight above 67.5 kg at mating had no positive effect on ovulation rate. The studies combined indicate that there is a positive relationship between ewe ovulation rate and liveweight, however, the relationship becomes less significant after reaching an optimum threshold weight, which is different for each breed.

A curvilinear relationship between fertility rate and a ewe's premating liveweight has been reported (Thomson et al., 2004; Kenyon et al., 2010; Aktaş et al., 2015; Corner-Thomas et al., 2015b). Both fertility and conception rates increase with increasing liveweight in commercial ewe-lamb flocks of up to 47.5 kg above which increases in liveweight resulted in no additional gains (Corner-Thomas et al., 2015a). Ewe lamb liveweight at mating has also been reported to be positively related to conception rate (McMillan and Moore, 1983) and lambing percentage (Dyrmundsson, 1973; Craig, 1982; Kenyon et al., 2004b). Liveweight is also positively related to litter size (Thomson et al., 2004; Kenyon et al., 2004b; Brown et al., 2005; Aktaş et al., 2015). Brown et al. (2005) and Ferguson et al. (2011) have demonstrated that when mated, heavier ewes tend to give more lambs per ewe.

Table 2.1 Summary of studies examining the relationship between liveweight and puberty onset, ovulation rate, conception rate and birth rate.

Reference	Animal details	Puberty onset	Ovulation rate per unit kg liveweight change	Ewe liveweight and Conception rate	Multiple birth rate/lambing percentage
Gunn and Doney (1975)			+ curvilinear, plateaus after 53 kg		
Allison AJ (1978)	2-tooth and older, Corriedale ewes			+	+ linear, twinning rate increased by 6% per 4.5 kg.
Morley et al. (1978)	5-year Corriedale ewes		+ linear, rate was 2% /1 kg difference in ewe live weight		
Meyer and French (	Finn-Romney cross	+	+	+	
Kelly and Johnstone (1982)			+ linear, rate was 1.6%/1 kg difference in ewe live weight		
McMillan and Moore (1983)		+	+	+	
Davis et al. (1987)	>= 1.5 years old Romney type		+, ewes with multiple ovulations at 1.5 years of age and at older age were heavier compared with ewes with one ovulation.		
Michels et al. (2000)	Mature merino		No association below 35-37.5 kg at mating. In heavier ewes 4%/ 1 kg within the rage 40-48 kg and 2%/1 kg increase in ewe live weight up to 53.5 kg		
Rutherford et al. (2003)	Mixed aged, Romney type ewes		+, at joining in small framed (2%) ewes but no significant in large framed (0.5%/1 kg) ewes. Overall, heavier ewes had greater ovulation rates compared with their lighter counterparts.		
Kenyon et al. (2004b)	3–5-year Romney & 2 tooth Romney composites		+, plateaus after 48.5 kg in composite Romney and at 58.7 kg in Romney		
Thomson et al. (2004)	2-tooth and older, Romney cross ewes				+ linear, lambing percentage increase of 1 % per kg of ewe live weight
Thompson and Oldham (2004)				+ linear, 3 foetuses/100 ewes joined	

Brown et al. (2005)				+ linear, lambing percentage increased by 0.2 per kg of ewe live weight
Kleemann and Walker (2005)	Merino	+ linear, rate was 1.8% /1 kg difference in ewe live weight		
Scaramuzzi et al. (2006)			+	
Ferguson et al. (2011)	2.5–3.5 years old Merino ewes		+ linear, 1.7 to 2.4 foetuses per 100 lambs joined	
Aktaş et al. (2015)	Central Anatolian Merino			+
Corner-Thomas et al. (2015a)	Composite (Romney type)		+linear from <32.5 to 47.5–52.4 kg	
Gabr et al. (2016)	2–8 years old Iranian Afshari ewes			+

+ indicates positive relationship; – : negative relationship; blank space: not indicated.

2.2.2 Lamb birth, survival, growth and weaning

The period between birth and weaning is very critical in sheep production, given that both the stock for replacement and sale are selected from the same mob of lambs produced in a season. During that time, growth and survival rates are essential selection criterion and are monitored over the season. The relationship between ewe liveweight and lamb birth weight, growth, survival and weaning weight has been extensively studied (Kelly and Johnstone, 1982; Kenyon et al., 2004a; Oldham et al., 2011; Thompson et al., 2011; Schreurs et al., 2012; Corner-Thomas et al., 2015a). A summary of studies on the effect of ewe liveweight and liveweight change on lamb growth, survival and weaning are given in Table 2.2.

Liveweight of the ewe at mating and liveweight change during pregnancy have been used to predict birthweight of the lamb in Australian studies (Oldham *et al*. 2011). Heavier ewes tend to give birth to lambs with heavier birth weights, which grow faster than low birth weight lambs, and are more efficient energy utilizers for tissue deposition (Kelly et al., 1996; Kenyon et al., 2004b; Oldham et al., 2011; Thompson et al., 2011; Schreurs et al., 2012; Behrendt et al., 2019; Hocking et al., 2019). Further, progeny growth rates are correlated with changes in maternal liveweight during pregnancy (Kenyon et al., 2004b; Morel et al., 2009). Progeny of ewes that are heavier at mating or have increased maternal weight at pregnancy grow faster to weaning (Greenwood et al., 1998; Kenyon et al., 2004b). However, Oldham et al. (2011) and Schreurs et al. (2012) in a meta-analysis of several studies, reported that ewe liveweight and liveweight change during gestation appear to give varying responses on the lamb birthweight, lamb weaning weight.

Lamb survival is also affected by ewe liveweight pre-mating and throughout pregnancy (Brown et al., 2005; Morel et al., 2009; Hocking. et al., 2011; Oldham et al., 2011; Aktaş et al., 2015). All studies suggest that heavier ewes at joining tend to have progeny with greater survival rates. Further, for a one unit liveweight gain during ewe pregnancy, lamb survival has been reported to increase by 0.38 % (Morel et al., 2009). However, Oldham et al. (2011) found little influence of liveweight and weight change during gestation on lamb birth and survival. Literature suggests that ewe liveweights can be managed to increase lamb survival, and this should be possible through nutritional management from a cost-benefit point of view (Morley et al., 1978; Rowe, 2003). Further, lamb birth weight plays a pivotal role in its perinatal lamb survival (Morley et al., 1978; Rowe, 2003). Optimum lamb birth liveweights range between 4 to 6.5 kg, and either below or above this range results in increased mortality (Greenwood et al., 1998; Greenwood et al., 2010; Hatcher et al., 2010; Thompson et al., 2011).

Table 2.2 Summary of studies examining the relationship between liveweight and lamb birth weight, growth, survival, weaning weight.

Reference	Animal details	Lamb birth weight	Lamb growth	Lamb survival	Lamb weaning weight
Hinch et al. (1985)	Booroola-Merino crossbred, on research station			Quadratic association with lamb birth weight declining at birth weight extremes.	
Holst et al. (2002)	Merino and crossbreeds, on research station.			(+), at less than 3 kg, no effect between 3 to 6 kg and (−) beyond 6 kg lamb birth for twin and triplet	
Thomson et al. (2004)	Romney crossbreeds under commercial conditions.			(+), at less than 3 kg, no effect between 3 to 9 kg and (−) beyond 9 kg lamb birth	
Brown et al. (2005)	Australian and New Zealand meat sheep and dual-purpose studs records.	0.012 kg per extra ewe liveweight gain pre-mating.			0.106 kg per extra ewe liveweight gain pre-mating
Casellas et al. (2007)	Ripollesa lambs under semi-intensive management.			Quadratic association with lamb birth weight declining at birth weight extremes.	
Morel et al. (2009)	Romney crossbreeds under commercial conditions.	+, 0.1 kg per 4.4 kg increase in ewe liveweight pre-lambing.	+, ADG of 0.001 kg/day per kg ewe liveweight gain pre-lambing	+, increased by 0.38 % per kg extra ewe liveweight gain during pregnancy	
Hatcher et al. (2009)	Merino sheep under commercial conditions.			Quadratic association with lamb birth weight declining at birth weight extremes.	
Van Der Linden et al. (2009)	Romney under commercial conditions.	+			
Oldham et al. (2011)	Wool merino ewes under different feeding levels (800, 1100, 1400, 2000 and >3000 kg DM/ha).	an extra 10 kg of ewe liveweight at joining increased lamb birthweight by approx 0.25 kg. A loss of 10 kg in ewe liveweight between joining and Day 100 of pregnancy reduced lamb birthweight by		Increased by 0.5% per extra kg of ewe liveweight at joining for lambs with low birthweight assuming maintenance of liveweight during pregnancy, 1.2% to Day 100 of pregnancy and 1.7% during late pregnancy.	

		approx. 0.33 kg, whereas gaining 10 kg from Day 100 to lambing increased birthweight by approx. 0.45 kg.		+ with lamb birth weight increased up to a birthweight of 4.5 kg and declined for single lambs weighing > 6.5 kg at birth.	
Greenwood et al. (2010)		Increase in birth weight by 0.03, 0.03 and 0.05 kg per extra ewe liveweight gain pre-mating, mating to 90 days and 90 days to lambing.			
Schreurs et al. (2012)	Romney ewes	+ with single birth but (−) with multiple birth.			
Aktaş et al. (2015)	Central Anatolian Merino sheep on-farm	+, with ewe liveweight pre-mating.		+ with ewe liveweight pre-mating.	+ with ewe liveweight pre-mating
Hocking et al. (2019)	Merino and Border Leicester crossbreeds raised on a research station.			Quadratic association with lamb birth weight declining at birth weight extremes.	+, a 10 kg higher ewe liveweight at conception resulted in 2.3 to 0.24 kg increase in lamb weaning weight. A 10 kg increase in early pregnancy weight resulted in a 2.4 to 0.47 kg and in late pregnancy resulted in 1.6 to 0.54 kg.
Behrendt et al. (2019)	Composite breeds under commercial conditions			Linear and quadratic association with lamb birth weight declining at birth weight extremes.	+, 10 kg change in ewe liveweight from joining to Day 90 resulted in a 1.8 to 2.0 kg difference in weaning weight.

+: positive relationship; −: negative relationship and blank space: not indicated.

2.2.3 Summary of liveweight relationships

The literature above suggests that liveweight plays a pivotal role in determining the outcome of production and reproduction traits and thus, underpins the importance of liveweight in sheep productivity. Liveweight affects puberty onset, fertility rates, pregnancy rate, fecundity, lamb growth and survival, all of which are critical in the sheep production cycle. It appears that there is a “minimal” or optimal range of liveweights for the best performance. Those threshold liveweight values can be used for decision making concerning selection for breeding and efficient resource allocation. It is thus imperative that farmers can accurately measure liveweight.

2.3 Liveweight measurement technique in sheep

The record of individual sheep performance can allow for differential management of sheep based on their respective liveweight change (Richards et al., 2006; Brown et al., 2014). There are several individual or collective methods of obtaining liveweight information of sheep. These methods range from the less efficient visual assessments (Suiter, 1994), laborious static manual balances, predictive body measurements as a proxy for liveweight, growth models, static electronic balances, walk-over balances and recently stereo imaging (Wilson, 2014; Brown et al., 2015). In commercial and research settings, conventional static weighing systems remain the principal technique of collecting liveweight information of sheep either individually or collectively (Brown et al., 2015). This review, therefore, will concentrate on the static electronic balance. The process of liveweight determination has made significant strides from manual recording to highly efficient automated balances.

Electronic weigh scales (Figure 2.1) have revolutionized liveweight measurement. These types of weighing scales can be managed automatically and can read the liveweight autonomously compared with manual weigh sales. Furthermore, these automated scales can be equipped with radio frequency identification (RFID) capacity and can store thousands of individual records. Consequently, automated electronic systems can produce liveweight data with minimal recording error. In combination, electronic identification and modern weigh systems allows individual lifetime data to be collected, thus, improving management outcomes. The usefulness of that data is dependent on consistent liveweights being collected over time. Currently, there are two common systems; static and walk-over weighing. Using the electronic scales, measurement efficiency is increased, whether it is placed within a confined area (Figure 2.1a) or strategically placed in the paddock for the animal to traverse over (Figure 2.1b) as part of their daily routine (Brown et al., 2015).

(a) (b)

Figure 2.1 Current weighing systems used to collect sheep liveweight data. (a) Static weigh scale and (b) Walk over weigh scale

The electronic weighing scale uses two methods to estimate an animal’s weight; (i) a measurement is taken when stability in animal movement (static) is detected or (ii) a measurement is determined by using a statistical process in which several readings taken by the processor are averaged over time (Smith and Turner, 1974). The first method is suitable for docile and restrained animals. It can be affected by fluctuations resulting from frequent movements in agitated and nervous animals leading to inaccuracies. The second method circumvents the challenges of the first method and therefore, it is more useful in field conditions with an accuracy of +1% achievable (Smith and Turner, 1974; Brown et al., 2015).

The RFID system is composed of three major components; an electronic tag on the animal, the RFID tag reader which links data to a transponder and a data processing unit (Richards et al., 2006; Geenty et al., 2007; Lee et al., 2008). Sheep getting weighed, have their liveweight recorded against a unique individual number, resulting in a RFID-linked weight record that can allow the liveweight of individual sheep to be tracked over time (Wilson, 2014; Brown et al., 2015).

2.4 Error in sheep liveweight measurement

In research and commercial livestock production, liveweight data can be used to make comparisons between liveweights at different time points, both within and between animals and groups (Wishart et al., 2017). To be able to generate consistent and comparable liveweights, the variation and error associated with these data need to be identified and controlled. Error can be defined as the difference between the “true” and the observed value (Drosg, 2009), arising from random or systematic effects (Taylor and Kuyatt, 1994; Bich et al., 2012). Whereas random error cannot be controlled, systematic error can be minimised. In theory, if this was achieved, then random effects would contribute all the unexplained variability in the measurement. Therefore, this review will concentrate on systematic error. Errors may arise from data collection, data recording, and computation of results (Elwood, 2017). Bich et al. (2012) listed a catalog of possible sources of error during measurement including incomplete definition of the measurement, imperfect realization of the definition of the measurement, nonrepresentative sampling, inadequate knowledge of the effects of environmental conditions on the measurement or imperfect measurement of environmental conditions, personal bias in reading instruments, finite instrument resolution or discrimination threshold, inexact values of measurement standards and reference materials, inexact values of constants and other parameters obtained from external sources and used in the data-reduction algorithms, approximations and assumptions incorporated in the measurement method and procedure, and variations in repeated observations of the measurement under apparently identical conditions. Wilson (2014) summarised the errors sources as either measurement (human and scale error) or animal related liveweight error. Furthermore, biological processes that are dynamic and can vary over time due to factors such as growth, physiological state, diurnal and seasonal variation, may cause within-subject variability (Kenyon et al., 2014; Brown et al., 2015; Wishart et al., 2017).

For indirect liveweight determination methods from N other independent variables through a functional relationship, the independent variable is assumed to be measured without error and that all error is attributed to the measurement of the dependent variable (Poole and O'Farrell, 1971; Greene, 2003; Alexopoulos, 2010; Bich et al., 2012; Dosne et al., 2016). However, when the independent variable is measured with error, this may lead to an alteration of the association between the outcome and the observed change in the independent variable (Cain et al., 1992; Bich et al., 2012). When the measurements of the independent variable are not exact, estimation based on the standard assumption leads to inconsistent parameter estimates even in very large samples (Hausman et al., 1995; Hausman, 2001; Pischke, 2007). It is therefore, imperative that only prediction models with minimum error rates or greater accuracy be used (Efron, 1983; Tibshirani and Tibshirani, 2009). Several measures of prediction model accuracy have been described (Moriasi et al., 2007; Li, 2017; Botchkarev, 2019). Alexander et al. (2015) suggested that mean absolute error percent or root mean square error percent of a prediction equation, should be less than 10% of the range of target or actual values. The following section will discuss potential causes of error with liveweight measurements.

2.4.1 Human error

Liveweight data obtained under field conditions are subject to an array of estimation biases. In the past weighing and recording liveweights were somewhat separate processes that were both manual and labour intensive. Individual tag numbers and animal's liveweight reading were recorded on paper by the operator (Wilson, 2014). Data were further entered into spreadsheets, validated, and analysed. The manual process, therefore, relied on the ability of the operator to accurately record information while operating the weighing apparatus (Collins and Atwood, 1981; Wilson, 2014). In a study to examine the presence and potential influence of these apparent investigator biases associated with spring-balance, Collins and Atwood (1981), reported that one out of the 11 participants had significantly different results and that 1.7% of the errors were due to the misreading of the scale. With the advent of automated electronic identification and scales, the potential for human error has been greatly reduced.

2.4.2 Technique and machine related error

Scale error results, are defined as dissimilar values of measurements obtained using different machines or when there is variation (spatial) in the results from the same machine (temporal). Lee et al. (2008) reported varying liveweight repeatability between static (0.99) and walk over weigh systems (crude, 0.35; crate base, 0.90, walk over base, 0.91). Galwey et al. (2013) reported that multiple weight recordings increased the accuracy of weight estimates in sheep. However, Bean (1946) observed that the use of a three day mean weight in swine introduced further error (2.1%) into the results instead of minimising it. Similarly, Bean (1948) and Wilson (2014) reported that a single weight in sheep was as reliable as the average of three consecutive daily weights, a conclusion supported by (Baker et al., 1947) in calves when uniform conditions were maintained.

2.4.3 Skeletal size, length, and fleece weight

Liveweight is not a good indicator of condition due to skeletal and frame size variations. A mature animal with medium fatness and with average liveweight, would weigh less when extremely thin. However, the same animal, when extremely fat would weigh more. Hammack and Gill (2001), suggested that if liveweight is to be used as an accurate measure of size, it must consider body condition. In addition, Brown et al. (2015) stated that, liveweights should be considered relative to the breed's mature average liveweight, and the animal's recent reference liveweight. Skeletal size increases with age as does liveweight but the rate of increase steadily decreases until mature liveweight is achieved (Wiener, 1967; Ho et al., 1989). In European and

Australian sheep, age at maturity varies from 25 to 50 months (Smith, 1956; Wiener, 1967; Cake et al., 2006).

The association of fleece weight on liveweight can depend on the sheep's age (Gonzalez et al., 1997), breed (Elliott et al., 1978; Gonzalez et al., 1997), and the season (Story and Ross, 1960). Elliott et al. (1978) reported that wool per unit liveweight for Coopworth, Perendale, and Cheviot were 10%, 18% and 39% respectively, which is less than Romney ewes. For seasonal effects, the rate of wool growth in Romney crossbred ewes varies considerably during the year, being highest in the summer and lowest in the late winter-early spring period (Story and Ross, 1960; Sumner et al., 1994). Cottle and Pacheco (2017) have also reported seasonality in the growth of wool of Romney sheep, with maximum wool length of 150 mm for single shorn and 75 mm for sheep shorn twice in a year (Table 2.3). Sheep with longer fleece and those in a wet environment weigh more than those with trimmed fleece (Story and Ross, 1960; Wiener, 1967; Elliott et al., 1978) and those in a dry environment. The contribution of fleece weight to the liveweight of the animal, can be accounted for through the use of liveweight-adjustment equations, while avoiding weighing sheep during or immediately after rain events generally negates the issue of fleece moisture content affecting liveweight (Brown et al 2015). However, those equations are not being used or adjusted for in electronic systems.

Table 2.3 Estimated insulation of the fleece (°C m2 d/MJ) in each month of the year for different shearing months of Romneys, assuming a shorn fleece length of 150 mm, seasonality amplitude of 19% of the mean, radius of 120mm, coat insulation of 0.141°Cm2/MJ/mm and rain and wind velocity (average 6.3 mm/rainy day and 13.1 km/h respectively).

	Insulation (°C m2 d/MJ)											
Shearing month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
January	1.12	2.08	2.90	3.78	4.38	4.86	5.49	5.86	5.96	6.44	6.91	7.35
February	7.53	1.10	2.01	2.91	3.56	4.10	4.74	5.17	5.35	5.86	6.37	6.92
March	7.12	7.61	1.06	2.00	2.72	3.31	3.98	4.47	4.73	5.28	5.82	6.40
April	6.63	7.23	7.65	1.08	1.87	2.53	3.23	3.77	4.11	4.70	5.29	5.90
May	6.17	6.80	7.32	8.07	1.03	1.76	2.49	3.10	3.52	4.15	4.78	5.41
June	5.73	6.38	6.92	7,76	8.06	1.00	1.77	2.44	2.95	3.62	4.28	4.95
July	5.29	5.96	6.53	7.37	7.78	7.97	1.03	1.77	2.37	3.08	3.79	4.48
August	4.83	5.52	6.11	6.95	7.38	7.69	8.16	1.03	1.74	2.50	3.26	3.99
September	4.30	5.03	5.64	6.49	6.94	7.27	7.84	7.98	1.01	1.84	2.65	3.42
October	3.68	4.45	5.10	5.95	6.42	6.78	7.37	7.62	7.44	1.05	1.94	2.76
November	2.95	3.77	4.46	5.32	5.82	6.22	6.81	7.10	7.06	7.40	1.07	1.96
December	2.09	2.98	3.73	4.59	5.14	5.57	6.18	6.51	6.54	6.98	7.33	1.02

Adapted from Cottle and Pacheco (2017)

2.4.4 Pregnancy and lactation

Pregnancy, especially foetus weight affects the gross weight of a ewe. The contribution by the foetus increases as the pregnancy progresses to maturity. Liveweight increases with increasing pregnancy-rank or number of foetuses carried by the ewe (Rattray et al., 1974; Russel, 1984). Conceptus and uterine weights have been reported to vary at Day 70 of pregnancy and near term in twin-bearing ewes (1.8−2.5 kg and 14−16.9 kg, respectively; (Rattray et al., 1974; Kenyon et al., 2007b). Conceptus weight can be accounted for through the use of generic or customised liveweight-adjustment equations (Wheeler et al., 1971; Freer et al., 1997; Brown et al., 2015; Ridler et al., 2017). The equations can vary by stage of pregnancy and plane of nutrition (Wheeler et al., 1971; Freer et al., 1997).

Ewe liveweight declines in early lactation then increases in late lactation. The decline in liveweight can be affected by the plane of nutrition, being highest in ewes on low and lowest in those on high (Peart, 1970; van der Linden et al., 2010). Lactation is the stage of highest nutrient requirement in the ewe's annual production cycle. Restrictions of nutrient intake in lactating ewes may result in the loss of body weight and body reserves of the ewe (Peart, 1982). The situation can be exacerbated by the number of lambs being reared, with liveweight losses being highest in multiple than single-bearing ewes.

2.4.5 Gut-fill variations and passage rates

Liveweight is a measure of the total body mass and includes muscle, fat, bone, organs and body fluids, gut-fill and fibre (Wishart et al., 2017). Liveweight is relatively stable over a short time period, although alters over time in response to environmental and physiological conditions (Coates and Penning, 2000b; Wishart et al., 2017).

The contents of the rumen (fluid and feed) can account for between 10 and 23% of total liveweight in ruminants (Hughes, 1976; Kingenberg, 2003; Moyo and Nsahlai, 2018). Liveweight fluctuations due to gut-fill in ruminants are known to be affected by time since last meal, feed and water consumption, age and size of the animal, time of day relative to sunrise, ambient temperature, and differences in grazing behaviour (Hughes and Harker, 1950; Whiteman et al., 1954; Hughes, 1976; Gregorini, 2012; Wilson, 2014; Wilson et al., 2015; Wishart et al., 2017).

During gestation and lactation, animals undergo structural and functional changes. Behavioural changes such as increased or decreased water intake, and gain or loss of appetite may be observed during these periods (Foot and Russel, 1979; Little et al., 1980; Kischel et al., 2017). Digesta's rate of passage through the rumen could also be altered by these changes. Rueda et al. (1990) showed that rates of particulate and liquid passage through the rumen were faster for pregnant than non-pregnant animals, higher in lactating animals than their non-lactating counterparts, however lower during the late than the early stages in gestation. Similarly, Hanks et al. (1993) working on beef cattle reported that particulate passage rate was greater for pregnant than non-pregnant cows. In pregnancy, the gut space is reduced by the growing foetus. A negative relationship between rumen volume and uterus volume in pregnant ewes has been reported (Forbes, 1969). Forbes (1969) reported that at day 72 of the gestation period the uterus and rumen volumes were 4.1 ± 0.7 litres and 6.6 ± 0.9 litres respectively, however by day 144 they were 7.7 ± 0.4 litres and 3.8 ± 0.4 litres respectively. In contrast, non-pregnant ewes had rumen volumes of 9.2 ± 0.6 litres throughout the same period. With reduced volume, animals in gestation period increase intakes and increase digesta retention time which results in reduced passage rate. These studies indicate that the variability in liveweight fluctuations seems to be influenced in a multifactorial way.

Several strategies can be used to reduce liveweight variation due to gut-fill. This includes removal of feed and water for fixed periods of time prior to weighing, standardizing weighing procedures, taking an average of multiple liveweights in a day or across a number of successive days, weighing at a specific time relative to sunrise, standardizing the feed offered before weighing and increasing the number of animals and repetitions of the study (Shrestha et al., 1991; Coates and Penning, 2000b; Burnham et al., 2009; Wilson et al., 2015; Wishart et al., 2017). Implementing such methodologies to reduce variation can, however, costly, be time-consuming and therefore not generally utilized except in experimental situations. It is also possible that farmers collecting liveweights are oblivious to the possible variability of the data.

Several studies have, however, demonstrated that the accuracy of liveweight measurement can vary because of differences of gut-fill and differential loss in gut-fill (Table 2.4) due to loss of ruminal content through faecal matter weight in each fasting period. It implies that all factors influencing gut-fill rate of passage should be investigated if accurate adjustment equations are to be generated. To date, technology companies have not yet incorporated weighing methodology in their systems to deal with this variation.

Table 2.4 Summary of studies examining the relationship between liveweight loss and nutrition, breed, age, reproductive status and time in pregnancy, and time off feed in ewes

Reference	Breed	Feeding level	Age	Reproductive status	Time in pregnancy (days)	Total time held (hours)	Weight loss in kg or % of initial weight	Liveweight loss prediction equation
Hughes (1976)			Lambs	Non pregnant		42–56	0.7–2 kg after eight hours; 1.0–2.0 kg after 12 hours	
			Two tooth	Non pregnant		36	1.5 to 3.3 kg after eight hours and 2.0 – 4.0 kg after 12hours	
Burnham et al. (2009)	Romney	Ryegrass/white clover pasture	10 months ewe lambs			24	After 2four hours ewe lambs	Y=0.01−0.014T+0.0007T2−0.000016T3, R^2=0.94
			Mature ewes	70 days in pregnancy	70	24	had lost 25.1%, ewes at day 70	Y=0.01−0.012T+0.0006T2−0.000011T3, R^2=0.72
			Mature ewes	130 days in pregnancy	130	24	lost 9.8% while ewes at day 130	Y=0.01−0.007T+0.0003T2−0.000004T3, R^2=0.72
			Mature ewes	Single-bearing, 130 days in pregnancy	130	24	lost 7.5%	Y=0.01−0.007T+0.0003T2−0.000004T3, R^2=0.82
			Mature ewes	Twin-bearing, 130 days in pregnancy	130	24		Y=0.01−0.006T+0.0003T2−0.000004T3, R^2=0.82
Wilson et al. (2015)	Coopworths	Ryegrass/white clover pasture removed from feed 2 hours before weighing	Mixed age ewes	Pregnant		24	After 12 hours they lost 6% and up to 7.9% after 20 hours	$Y=0.0077x + 0.0002x^2$, (R^2=0.9794)
	Coopworths	kept in the yards after shearing and fed baleage prior to the experiment	Mixed age ewes	Pregnant		24	After 12 hours they lost 1.5 % and up to 2.0 after 20 hours	$Y=0.0018x + 2E{-}05x^2$, (R^2=0.9954)

Wishart et al. (2017)	Scottish blackface, Lleyn ewes, their crossbred lambs	Improved pasture	1.5, 2.5, 3.5 &4.5 years	Non pregnant, non-lactating	6	after 3 hours, 3.5%; after 6 hours 5.6%

2.4.6 *Species, Grazing and diurnal variation*

Ruminants have different feeding habits depending on whether they are grazers (cattle, sheep) or browsers (goats). The differences in the diets and processes associated with feeding behaviour among these classes of animals can effect on rates of passage of liquid and solid phases in the rumen and their rumen fill (Lechner et al., 2009; Moyo and Nsahlai, 2018). Sheep have lower mean retention times of solid in the rumen than cattle (Lechner-Doll et al., 1991; Bartocci et al., 1997). Parra (1978) also demonstrated that there were higher passage rates for lighter herbivores than larger herbivores with diet quality held constant.

The pattern of grazing events dictates how much an animal ingests within a given time and thus influences the liveweight and liveweight change throughout the day. Grazing strategy differences are known to affect passage rates and rumen fill levels in cattle (Oshita et al., 2008). Ruminants have three to five grazing events every day, with the greatest intake periods being early in the morning and in the late afternoon (Gregorini et al., 2008; Gregorini, 2012). Rook and Penning (1991) reported that 70-99% of grazing occurs during daylight, with 25-48% occurring in the four hours prior to sunset. In cattle, close to one third of their total grazing time occurs during dawn (Gregorini et al., 2008; Hilario et al., 2017). Therefore, the time at which an animal is weighed can affect the amount of gut-fill and thus its liveweight.

Grazing patterns vary between animals depending on quality and type of herbage and the environment (Kirby and Stuth, 1982; Ginane and Dumont, 2010; Lin et al., 2011). Orr et al. (1997) stated that herbage DM % varied during a 12-hour period between 15-24% grass and 12-18% clover with the most significant change happening from morning to noon. Starch content also changed from 3.0-4.1% and 3.6-8.7%, for grass and clover, respectively. This further indicates the time of day can influence liveweight.

According to Hamilton et al. (1995), the greatest diurnal variation in estimated liveweight was observed between 11am-1pm and the lowest variation reported at 9 am and 4 pm, with sunrise at 6 am. Small ruminants are selective feeders (Ginane and Dumont, 2010; Lin et al., 2011) and are more inclined to feed on lower dry matter and lower starch pastures which are more easily digestible. It is likely that liveweight is overestimated when an animal is fed on low dry matter, lower starch, and high concentrate diets.

2.4.7 *Seasonal ambient temperature variations*

Seasonal temperature variations affect feed and water intake and consequently on the digesta passage rate, and thus liveweight loss. Animals kept in cold environment consume more feed, have increased digesta flow rate but grow slower as more energy is converted into heat to maintain their body temperature (Young, 1981). It appears that dry matter digestibility decreases during winter conditions (Christopherson, 1976) which can be associated with increased gut motility, passage rate and circulating thyroid hormone. (Kennedy et al., 1976) demonstrated in a trial with sheep that the flow of dry matter and organic matter was greater through the abomasum during the cold exposure, −1.0° to 1.0°C, than during the warm exposure, 18° − 21°. It was also noted that when the ambient temperature was lowered from 21 to 0^{o}C, the mean retention time of solid digesta in the rumen decreased by 20% (Kennedy, 1985). Heat exposure effects counter to those of the cold exposure. Bernabucci et al. (1999) and Miaron and Christopherson (1992) working on heifer and steer trials, respectively, reported that the rumen outflow rate was lower at higher temperatures. There was also reduced dry matter intake and increased the water intake compared with the period under the thermal comfort zone. True liveweight seems to be overestimated in cold exposure and underestimated during the hot exposures. These data further indicate that time of day when the measurement occurs can influence liveweight and liveweight loss.

2.5 Body condition score in sheep

Body condition score (BCS) is a subjective measure which provides estimates of body condition for farmers and technicians to describe energy reserve levels under practical production conditions (Morris et al., 2002; Vieira et al., 2015). BCS circumvents the shortcomings of using liveweight alone to predict body condition. It is easily learned, cost-effective and requires no specialised equipment (Kenyon et al., 2014; Brown et al., 2015; McHugh et al., 2019). In addition, Jefferies (1961) suggested that BCS in sheep could be used to allocate feed efficiently, detect subtle changes in the body condition not noticeable by visual inspection, allow farmers to be more aware of major losses in body condition such as wasting and to be able to follow trends in nutrition and liveweight. The original purpose of the technique was four-fold and included; (1) control of condition/nutrition of sheep, for more efficient utilization of available food supplies; (2) detection of small differences in body condition not noticeable by outside appearance; (3) empowerment of farmers to be immediately aware of major losses in body condition; (4) monitoring of trends in nutrition and liveweight. BCS is thus considered a useful way for farmers to monitor the condition of their flock and estimate the required plane of nutrition (Kenyon et al., 2014).

2.6 The effect of BCS in sheep productivity

Body condition score is an indicator of the energy balance of a ewe which is an important factor in determining the number and weight of lambs weaned (Scaramuzzi et al., 2006; Kenyon et al., 2014). Therefore, it might be expected that ewes of lower BCS will display reduced reproductive performance in comparison with those of greater BCS (Kenyon et al., 2014). Several authors have reported a positive relationship between BCS and reproductive traits (Tables 2.5 - 2.7). The following sections give a brief summary of the known relationships between BCS and sheep productivity.

2.6.1 Breeding season, Ovulation rate and conception rates

Body condition score is positively associated with breeding season, ovulation rate and conception rates (Kenyon et al., 2014). Table 2.5 gives a summary of the known relationship between BCS and reproductive traits from breeding to pregnancy. The relationships, however, seems to be confined to specific BCS ranges and can be affected by breed differences (Gunn and Doney, 1979; Gunn et al., 1991; Gunn et al., 1991a; Kenyon et al., 2014). Body condition score is also positively related to conception rate within certain BCS ranges above which the relationship changes (Gunn et al., 1991a; Sejian et al., 2010; Kenyon et al., 2014). At BCS between 2.5 and 3.5, the relationship with fertility and pregnancy rates plateau.

2.6.2 Number of foetuses, Number of lambs born and lamb survival

The relationship between BCS and number of foetuses, number of lambs born and lamb survival is established (Gunn et al., 1969; Adalsteinsson, 1979; Kleemann and Walker, 2005; Abdel-Mageed, 2009; Kenyon et al., 2014). The results indicate that generally, the relationship is positive although it is affected by breed differences (Kleemann & Walker 2005; Gunn et al. 1998, 1991a).

Most authors (Table 2.5) have reported a positive relationship between BCS and the number of lambs born per ewe (Gunn et al., 1969; Adalsteinsson, 1979; Kleemann and Walker, 2005; Abdel-Mageed, 2009; Aliyari et al., 2012; Kenyon et al., 2014). In contrast, McInnes and Smith (1966), Geisler and Fenlon (1979) and (Rozeboom et al., 2007), all reported that the number of lambs born per ewe is independent of ewe BCS in Merino. The observed variation between studies could be attributed to breed differences as well as the possibility that the positive relationship between BCS and the number of lambs born may not be linear and instead curvilinear. At body condition score between 2.5 and 3.5 the relationship appears to plateau and later decline.

Body Condition Score has been reported to have either no effect on lamb survival to weaning (Al-Sabbagh et al., 1995; Oldham et al., 2011) or a positive effect (Litherland et al., 1999; Dodds and Everett-Hincks, 2008). Kleemann and Walker (2005) and Rozeboom et al. (2007), observed a positive curvilinear relationship between BCS and singleton lamb survival in Merino ewes, with a diminishing response as BCS increased above 3.0 but in twins, the relationship remained linear (Table 2.6). Lamb survival is a binomial trait and therefore, relatively large numbers of lambs are needed to be able to detect differences. Inadequate numbers may have contributed to the lack of effect observed in some lamb survival data.

2.6.3 Lamb birth, growth and weaning weight

The relationship between BCS and change in BCS, and lamb growth to weaning is also well studied (Thompson et al., 2011; Kenyon et al., 2014; Behrendt et al., 2019). Ewe BCS has been reported to have either no influence on lamb growth to weaning (Gibb and Treache, 1980; Litherland et al., 1999; Thompson et al., 2011) or weaning weight (Al-Sabbagh et al., 1995; Litherland et al., 1999; Aliyari et al., 2012; Verbeek et al., 2012), or a positive effect on lamb growth (Gibb and Treache, 1980; Kenyon et al., 2004a; Kenyon et al., 2011a; Mathias-Davis et al., 2013; Behrendt et al., 2019) and weaning weight (Molina et al., 1991; Sejian et al., 2010; Behrendt et al., 2019). Table 2.7 outlines a summary of studies examining the relationship between BCS and lamb growth, survival and weaning. The variation between studies may be due to differences in the timing of the BCS measurement, the levels of BCS being compared, the plane of nutrition, and the number of lambs born and reared per ewe. For those studies reporting a positive effect, the relationship appears to be observed at BCS range of 2.5 to 3.0.

As indicated previously, BCS and productivity for many sheep traits are positively related. However, at higher BCS there is a plateauing effect. This non-linear relationship means that the biggest gain can be achieved by reducing the number of ewes with the lowest BCS in a flock or ensuring that all individuals are above a target threshold. To manage an animal to its optimum BCS, it must be accurately and repeatedly measured.

Table 2.5 Summary of studies examining the relationship between BCS and breeding season, ovulation rate and conception rate

Reference	Breed	When BCS recorded and range tested	Nutritional treatment(s) during examination period[a]	Length of breeding season relationship	Ovulation rate relationship	Conception rate relationship
Gunn et al. (1969)	Scottish Blackface	Breeding, 1.5 and 3.0	Low, maintenance, high			
Gunn et al. (1972)	Scottish Blackface	Breeding, 1.5 and 3.0	Fed to maintain of BCS		+	
Gunn and Doney (1975)	Scottish Blackface	Breeding, 1.0 to 3.0	Low, maintenance, high		+	
Gunn and Doney (1979)	Cheviot	Breeding, 2.0 and 3.0	Fed to maintain of BCS		+	
Newton et al. (1980)	Masham	Breeding, 2.0 and 4.0	Fed to maintain of BCS	+ late in breeding season	+	
Knight (1980)	Romney	Pre-breeding	Commercial conditions		+	
Rhind et al. (1984a)	Scottish Blackface	Breeding, 1.8 and 2.8	Fed to maintain of BCS		+	
Rhind et al. (1984b)	Greyface	Pre-breeding, 2.5–3.0 and 3.25–3.75	Fed to maintain of BCS		+	
McNeilly et al. (1987)	Scottish Blackface	Pre-breeding, 1.8–2.9	Fed to maintain of BCS		+	
Gunn et al. (1988)	Beulah-Speckled-face & Brecknock Cheviot	Breeding, <=1.5,1.75–2.5,2.25–2.5 and >=2.75	Low, high		+ and + to 2.25–2.5 in two differing breeds	
Gunn et al. (1991a)	Welsh Mountain & Brecknock Cheviot	Pre-breeding, <=2.25, 2.5 and 2.75	Low, high		+ and + to 2.5 in two differing breeds	+ and + to BCS 2.5 in two differing breeds
Gunn et al. (1991a)	Welsh Mountain & Brecknock Cheviot	Pre-breeding, <=2.25, 2.5 and >=3.0	Low, maintenance		+ and + to BCS 2.5 – 2.75	+ and + to BCS 2.5 – 2.75
Forcada et al. (1992)	Rasa Aragonesa	Breeding, <=2.25 and 2.75	Fed to maintain of BCS	+	+	
Rondon et al. (1996)	Rasa Aragonesa	Breeding, <=2.5 and >=2.75	High	+		
Kleemann and Walker (2005)	Merino	Breeding	Commercial conditions		+	

Sejian et al. (2010)	Malpura	Pre-breeding, 2.5,3.0–3.5 and 4.0	Fed to maintain of BCS	+ to BCS 3.0–3.5 then −
Corner-Thomas et al. (2015a)	Romney type	Pre-breeding, 2.0, 2.5, 3.0, 3.5 and 4.0	Commercial conditions	+ to BCS 3.0 then NR 3.0–4.0

Adapted from Kenyon et al. (2014) and modified. [a]unless otherwise stated there are no interactions between nutritional treatments and BCS. N/S, noted stated; +, positive relationship; − negative relationship.

Table 2.6 Summary of studies examining the relationship between BCS and the number of embryos/foetuses, number of lambs born and lamb survival

Reference	Breed	When BCS recorded and range tested	Nutritional treatment(s) during examination period[a]	BCS and number of foetuses per ewe relationship	BCS and number of lambs born relationship	BCS and lamb survival relationship
Gunn et al. (1969)	Scottish Blackface	Breeding, 1.5 and 3.0	Low, maintenance, high		+	
Pollott and Kilkenny (1976)		Breeding, BCS range not stated	N/S		+	
Adalsteinsson (1979)	Icelandic	Breeding, 2.0 and 4.0	Commercial conditions		+ to BCS 3.0–3.5	
Newton et al. (1980)	Masham	Breeding, 2.0 to 4.0	Fed to maintain of BCS		+	
Gunn et al. (1988)	Beulah-Speckled-face & Brecknock Cheviot	Pre-breeding, <=2.25,2.5–2.75,>=3.0	Low, high		BCS 2.5–2.75 greater than BCS <=2.25 and >=3.0	
Rhind et al. (1984b)	Greyface	Breeding, 2.75, 3.0, 3.25,>=3.5; Pre-breeding, 2.5–3.0 and 3.25–3.75	NS Fed to maintain of BCS	–	–	
Gunn et al. (1988)	Beulah-Speckled-face & Brecknock Cheviot	Breeding, <=1.5,1.75–2.5,2.25–2.5 and >=2.75	Low, high	+ in one breed, NR in second breed		
Gunn et al. (1991a)	Welsh Mountain & Brecknock Cheviot	Pre-breeding, <=2.25, 2.5 and 2.75	Low, high	In high BCS + to 2.5, no effect low feeding		
Gunn et al. (1991)	Cheviot	Pre-breeding, <=2.25, 2.5 and >=3.0	Maintenance, high		BCS 2.5–2.75 greater than BCS <=2.25 and >=3.0	
Al-Sabbagh et al. (1995)	Norduz	Pre-lambing, BCS 2.5, 3.0, 3.5	High			NR
Gonzalez et al. (1997)	Merino & Corriedale	Breeding, 2.0, 2.5, 3.0, 3.5 and 4.0	Commercial conditions		+	

Litherland et al. (1999)		Pre-lambing, 1.5 and 2.5	Low, high			+ in one of two studies
Atti et al. (2001)	Fat-tailed Barbarine	Pre-breeding, BCS range not stated	Commercial conditions		+ to BCS 3.0–4.0	
Kenyon et al. (2004b)	Romney	Breeding, 1.5 to 4.0	Commercial conditions	+ to BCS 2.0 in one breed and + to BCS 3.0 in second breed		
Kleemann and Walker (2005)	Merino	Breeding, BCS range not stated	Commercial conditions	+	+	+
Rozeboom et al. (2007)		Pre-lambing, 1.5 to 3.5	N/S		NR	
Abdel-Mageed (2009)	Ossimi	Pre-breeding	Maintenance		+ to BCS 2.5 then − after for BCS 4.0	
Kenyon et al. (2011a)	Romney	Mid-pregnancy, <=2.0, 2.5 and >=3.0	Maintenance, high			BCS 2.5 lower than <=2.0
Oldham et al. (2011)	Merino	Day 100 of pregnancy, 2.0 and 3.0	Various feeding levels			NR
Kenyon et al. (2012a)	Romney type	Mid-pregnancy,2.0, 2.5 and 3.0	Medium, high			BCS 2.5 lower than 2.0
Aliyari et al. (2012)	Afshari	Pre-breeding, 2.0, 2.5, 3.0 and 3.5	Ad libitum		NR	
Corner-Thomas et al. (2015a)	Romney type	Pre-breeding, 2.0, 2.5, 3.0, 3.5 and 4.0	Commercial conditions	+ to BCS 3.5 then NR 3.5–4.0		

Adapted from Kenyon et al (2014). [a]Unless otherwise stated there are no interactions between nutritional treatments and BCS. NR, no relationship or effect; N/S, not stated; +, positive relationship; −, negative relationship.

Table 2.7 Summary of studies examining the relationship between BCS and lamb birth and weaning weight and lamb growth to weaning

Reference	Breed	When BCS recorded and range tested	Nutritional treatment(s) during examination period[a]	BCS and lamb birth weight relationship	BCS and lamb growth relationship	BCS and lamb weaning weight relationship
Gibb and Treache (1980)		Pre-breeding, 2.4 and3.2	Low, high	NR	+	
Gibb and Treacher (1982)		Day 90 pregnancy, 2.6 and 3.3	Low, high in pregnancy, high in lactation	NR	NR	
Molina et al. (1991)	Machega lambs	Pre-lambing, <2.5, 2.5–3.0, >3.0		+		+ to BCS>3.0
Al-Sabbagh et al. (1995)	Fat-tailed Barbarine	Pre-lambing, BCS 2.5, 3.0, 3.5	High	NR		NR
Atti et al. (2001)	Norduz	Pre-lambing, <2 and >3	Maintenance		+	
Litherland et al. (1999)		Pre-lambing, 1.5 and 2.5	Low, high		NR	NR
Kenyon et al. (2004a)	Romney	Breeding, 1.5 to 4.0	Commercial conditions	BCS 3.5–4.0 >3.0	+	
Sejian et al. (2010)	Malpura	Pre-breeding, 2.5, 3.0–3.5 and 4.0	Fed to maintain BCS	+		+ to BCS 3.0–3.5
Kenyon et al. (2011a)	Romney	Mid-pregnancy, <=2.0, 2.5 and >=3.0	Medium, high	NR		BCS <=2.0 lower than 2.5
Oldham et al. (2011)	Merino	Day 100 of pregnancy, 2.0 and 3.0	Various feeding levels	+ in two of four studies		
Kenyon et al. (2012a)	Romney	Mid-pregnancy,2.0, 2.5 and 3.0	Medium, high	NR		+ to BCS 2.5
Kenyon et al. (2012b)	Romney	Mid-pregnancy, 2.0, 2.5 and 3.0	Medium, high	NR		+ to BCS 2.5
Verbeek et al. (2012)		BCS mid pregnancy, 2.0, 2.9 and 3.7	Fed to maintain BCS	NR		NR
Aliyari et al. (2012)	Afshari	Pre-breeding, 2.0, 2.5, 3.0 and 3.5	Ad libitum	NR		NR
Behrendt et al. (2019)	Composite breeds under commercial conditions	Pre-mating, mating to pregnancy, pregnancy to lambing, 2.4–2.5, 2.8–3.0, 3.2–3.4 and 3.6–3.8	Fed to maintain BCS	+	+	+

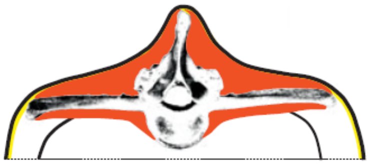
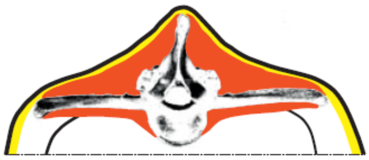
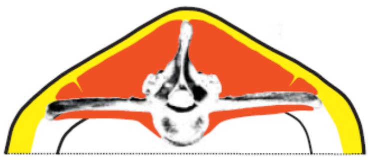
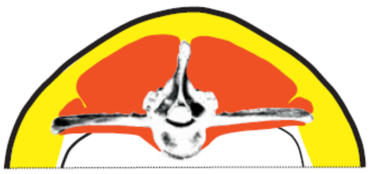
Hocking et al. (2019)	Merino and Border Leicester crossbreeds raised on a research station.	Pre-mating to pregnancy (50, 90 140 days), 2.5, 2.8, 3.2, 3.6	Fed to maintain BCS	+		+

Adapted with modifications from Kenyon et al (2014). [a]Unless otherwise stated there are no interactions between nutritional treatments and BCS.
NR, no relationship or effect; Blank space, not stated; +, positive relationship; −, negative relationship.

2.7 BCS techniques in sheep

The BCS of an animal is assessed by the palpation of the lumbar region, specifically on and around the backbone (spinous and transverse processes) in the loin area, immediately behind the last rib and above the kidneys to inspect the degree of fat and tissue coverage (Jefferies, 1961; Teixeira et al., 1989; Kenyon et al., 2014; Brown et al., 2015; Morel et al., 2016). First published by Jefferies (1961) in Scotland sheep, the technique was based on a 0 to 5 scale, including only whole units (Table 2.8). The scoring system was subsequently modified by Russel et al. (1969), working on English meat sheep who introduced the concept of 0.5 and 0.25 units. Different scales have been used to estimate BCS including; 0 to 5 (Russel et al., 1969; Russel, 1984; Sezenler et al., 2011); 1 to 5 (Thompson and Meyer, 1994; Kenyon et al., 2004a, 2004b; Morel et al., 2016) and a scale of 1 to 10 (Everitt, 1962; Sanson et al., 1993). The point intervals used in the studies has also differed; 0.5 and 0.25 (Russel, 1984; van Burgel et al., 2011).

Table 2.8 Description of the BCS technique and an illustration of the vertebra and ribs and approximate muscle and fat distribution.

Grade	Description	Illustration
Score 1	The spinous processes are prominent and sharp. The transverse processes are also sharp, with fingers passing easily under the end of this process. The eye muscle areas are shallow with little to no fat cover.	
Score 2	The spinous processes are smooth but still prominent. The individual processes can still be felt but only as fine corrugations. The transverse processes are smooth and rounded. However, it is still possible to pass the fingers under the ends of the processes with some pressure. The eye muscle areas are of moderate depth, but have sparse fat cover.	
Score 3	The spinous processes are smooth and rounded, and individual bones can only be felt with some pressure applied. The transverse processes are also smooth and are well covered. Firm pressure is required to feel over the ends. Eye muscle area is full and covered by a moderate degree of fat.	
Score 4	With pressure applied, the spinous processes can just be detected, although the ends of the transverse processes cannot. Eye muscle areas are full with a thick covering of fat.	
Score 5	Even with firm pressure applied, the spinous processes cannot be detected. Due to a high level of fat adjacent to the spinous process, a depression directly above where the spinous processes would normally be felt may be present. It is not possible to detect the transverse processes. The eye muscle areas are very full with very thick fat cover. It is possible to have significant deposits of fat cover over the rump and tail.	

Adapted from Kenyon et al. (2014).

2.8 Reliability of technique

Due to the subjective nature of BCS, its reliability within and between assessors could be a significant stumbling block to the potential use and effectiveness of this technique. Evidence from studies suggests that the repeatability of the BCS within assessor has varied from low to high (Table 2.9). Overall, the data suggest that inexperienced assessors can have difficulty achieving consistency between assessments (Everitt, 1962; Yates and Gleeson, 1975), whereas experienced assessors appear to be able to achieve high herbage levels of consistency, even when assessing ewes to 0.25 units (Teixeira et al., 1989; van Burgel et al., 2011; Phythian et al., 2012).

Body condition score technique has been demonstrated to exhibit high repeatability with up to between 80% and 90% within for experienced assessors (Teixeira et al., 1989). Body condition score techniques has, however, been reported to also have low (5 – 27%) between and 16 – 44% within repeatability for inexperienced assessors (Yates and Gleeson, 1975). Yates and Gleeson (1975), also stated that assessors found the later stages of pregnancy particularly difficult to assess. This may suggest that changes in the shape of a ewe in late pregnancy can influence the ability of the assessor to accurately determine BCS and may warrant investigation. All studies suggest that reliability and repeatability appear to be the primary limitations of BCS measurement.

Some guidelines have been suggested to improve consistency in BCS estimation; (i) the variation could be reduced by having two different assessors providing an estimate for each ewe (ii) use of 'condition score' models (score 1−5, in 0.5 units) such as those developed by 'Lifetime wool' to reduce between-operator bias, (iii) allowing for assessor calibration and training and (iv), a short period of recalibration of assessors (Evans, 1978; Calavas et al., 1998; Curnow et al., 2011; van Burgel et al., 2011; Phythian et al., 2012; Corner-Thomas et al., 2015b). Adherence to such guidelines can be, costly, time consuming and may require committed operators.

Table 2.9 Repeatability of the BCS technique between and within assessors

Reference	Scale system used and smallest units	Assessor experience level	Repeatability between assessors	Repeatability within assessors
Everitt (1962)	1–10, whole units	Inexperienced	Variation between and within assessor over time, values not stated	
Russel et al. (1969)	0–5, 0.25 units	Not stated	>70% absolute agreement, <20% differed by 0.5 BCS units and <10% by 1.0	>80%, <15% of observations varied by 0.5 BCS units and <5% by 1.0 units
Milligan & Broadbent (1974)	0–5, units not stated	Not stated		r = 0.49–0.67
Yates & Gleeson (1975)	0–5, 0.25 units	Inexperienced	r^a = 0.05–0.27	r = 0.16–0.44
Evans (1978)	0–5, 0.5 units	Not stated	r = 0.81	r = 0.88
Teixeira et al. (1989)	0–5, 0.25 units	Experienced	80%	90%
Calavas et al. (1998)	0–5, 0.25 units			k^b = 0.6–1.0
Shands et al. (2009)	0–5, 0.25, 0.5 units	Mixed skill level	r = 0.73–0.89	r = 0.64–0.84
van Burgel et al. (2011)	0–5, 0.25, 0.5 units	Experienced		0.2 unit difference in BCS mean of animals tested between assessments
Phythian et al. (2012)	0–5, 1.0 and 0.5 units	Experienced	k = 0.4–0.6	k = 0.6–0.7

r^a = correlation. k^b = weighted kappa analysis: <0.4 (poor level agreement); 0.4–0.75 (fair–good); >0.75 (excellent). Adapted from Kenyon et al. (2014)

2.9 Use of BCS by the sheep industry

Use of BCS to actively monitor body condition changes at key stages in production (i.e. mating, pregnancy diagnosis, lambing and weaning) is recommended as a cheap and cost-effective complement to liveweight measurement, as it circumvents the aforementioned limitations of the liveweight (van Burgel et al., 2011; Kenyon et al., 2014; Brown et al., 2015; Morel et al., 2016) . Despite the numerous advantages of using BCS over liveweight alone to better manage their flocks, it is uncommon for producers/farmers to regularly and objectively do so. A survey of sheep producers indicated that even though 96% reported monitoring the body condition of their sheep, only 7% of the producers did hands-on BCS assessment of ewe condition to estimate the energy requirements of their sheep (Jones et al., 2011). In New Zealand, a greater proportion of farmers who do hands-on BCS Corner-Thomas et al. (2016) reported that the proportion of farmers using BCS as a management tool at 40%. Combined these findings indicate that there is a sizable number of farmers not using BCS. Instead, most farmers relied on the visual inspection, a method that has been demonstrated to be inaccurate (Besier and Hopkins, 1989). The possible reasons for the low uptake of BCS are that, firstly, although the ideas in the guidelines appear plausible, they are somewhat unpractical and secondly, end users of the technique may be unaware of the guidelines (Kenyon et al., 2014). It is possible that if BCS can be accurately indirectly measured without the physical touching of the sheep, BCS use may be increased within the industry.

The use of BCS still remains low by the sheep farmers mainly due to its perceived arduous practical requirements. However, given BCS's numerous advantages (such as circumventing the effects of gut-fill, skeletal size, fleece weight and wetness and physiological state of sheep) over liveweight as a flock management tool, better technology (i.e. hands-free) is needed to increase its uptake.

2.10 Other methods for body condition assessment

2.10.1 Digital image analysis

Manual determination of body condition score can be labour intensive, requires a trained and experience hand and is not conducive to the frequent collection of data in an extensive commercial context. Digital image analysis offers an alternative method to continuously collect and automatically monitor body conformation measurements for BCS estimation BCS in real-time (Bell et al., 2018). Using digital cameras, animal images are taken from above the animal to relate body shape angels/curvatures around the hook bones and caudal area to the body condition. This method has been successfully used to estimate the BCS of cattle (Bewley et al., 2008; Azzaro et al., 2011) and may have potential for estimating the BCS of sheep in a paddock (Burke et al. 2004). The accuracy of the method is affected by the camera angle used to obtain the image, requires that images will need to have clearly defined boundaries that enable the measurement of certain truss points and curvatures given that colour uniformity and wool cover will likely distort such images (Burke et al. 2004). Given these challenges and the impracticality of strict restraint of animals for imaging, it is unlikely that the technology will be used by livestock managers to produce BCS estimates of sheep.

2.10.2 Ultrasound

Backfat thickness determined by ultrasonography along with the BCS can be used to assess the energetic and body state in a number of animal species (Zulu et al., 2001; Broring et al., 2003; Chay-Canul et al., 2016; Silva et al., 2016). In sheep, several studies have reported a positive correlation ($0.45 \leq r \leq 0.67$) between ultrasound measurements and body composition in wool and non-wool bred sheep (Junkuszew and Ringdorfer, 2005; Chay-Canul et al., 2016; Chay-Canul et al., 2019). Chay-Canul et al. (2019) stated that ultrasound measurements around the area of the *Longissimus dorsi* muscle (LDA) had stronger correlation with BCS than measurements at the thoracic region and was a better indicator of body reserves. All studies in sheep, expressed misgivings about the relative difficulty in taking ultrasound measurements in sheep due to the fleece cover and the looseness of the outer layer of subcutaneous fat. Ultrasound is a potential method for body condition assessment in sheep. However, it is hands-on, requires a knowledgeable operator and may not be appropriate for woollen sheep. Given the practical limitations above, it is also unlikely that the technology will be used by livestock managers to produce BCS estimates of sheep.

2.10.3 Liveweight and liveweight change

There are well established positive relationships between liveweight and BCS. Most authors have reported a linear relationship (Koycu et al., 2008; Kenyon et al., 2014; Morel et al., 2016) while Teixeira et al. (1989) suggested a curvilinear relationship (Table 2.10). The magnitude of the liveweight difference per unit BCS is affected by a number of factors including breed, age and physiological status of the animal (Frutos et al., 1997; Kenyon et al., 2004a, 2004b; Freer et al., 2007; Kenyon et al., 2014; McHugh et al., 2019). The variations in breeds are likely due to differences in frame size, conformation, standard reference weight and differences in fat distribution throughout the body (Geisler and Fenlon, 1979; Russel, 1984). Ho et al. (1989) stated that as animals grow, their frame size increases, until the bones cannot grow further, and this is their mature size. Ewes attain their mature liveweight between 25 to 50 months of age (Smith, 1956; Wiener, 1967; Cake et al., 2006), after which the relationship between the liveweights difference per unit BCS would be expected to be stable and more predictable. Consequently, it should be possible, although not yet known to predict the BCS using liveweight as a proxy. It has been stated that an additional unit of BCS equates to 3.3−11 kg in liveweight (Table 2.10). Due to the number of factors that can potentially affect kg/BCS unit change, it is difficult to use a simple standard weight change to predict BCS or BCS change. However, if a model could be developed to predict the relationship between weight and BCS, it would likely be used by farmers when they weigh sheep to estimate BCS.

Table 2.10 Average change required in liveweight per unit of BCS and nature of association across breeds and sheep classes

Reference	Timing of measurement	Breed	Sample size	Age (years)	Scale (decimal units)	Nature of Relationship[a]	Liveweight change	R^2
Jefferies (1961)							6.8	
Russel et al. (1969)		Scottish Blackface	273	Adult	0–5 (0.25,0.5)	linear	10.6	0.87
Geisler and Fenlon (1979)	Breeding	Eight breeds			0–5 (0.5)	linear	3.3–7.8	
Hossamo et al. (1986)	Breeding	Awassi					5.8	
Teixeira et al. (1989)	Dry	Rasa Aragonesa		10	0–5 (0.25)	Curvilinear	7.0–16	
Sanson et al. (1993)	Dry	Western-range	14	mature	1–9 (0.5)	Linear	5.1	0.78
Frutos et al. (1997)	Non pregnant, Non lactating	Churra	35	5–7		Linear	5.6	
Kenyon et al. (2004a)	Breeding /mating)	Romney	435	5.0	1.0–5.0 (0.5)	Linear	7.3	0.99
Kenyon et al. (2004b)	Breeding /mating)	Romney	1780	3–5	1.0–5.0 (0.5)	Linear	7.9	0.99
Kenyon et al. (2004b)	Breeding /mating)	Romney composite	692	3–5	1.0–5.0 (0.5)	Linear	4.8	0.99
Freer et al. (2007)	Dry	Polwarth X SA Merino	47	Adult	0–5	Linear	6.3	0.27
	Dry	Polwarth X SA Merino	60	maiden	0–5	Linear	7.3	0.28
	Dry	Saxon Merino	44	Adult	0–5	Linear	5.6	0.29
	Dry	Saxon Merino	42	maiden	0–5	Linear	7.0	0.31
	Lactating	South Aust Merino	10		0–5	Linear	5	0.28
	Lactating	Saxon Merino	10		0–5	Linear	5.5	0.16
			90, 58		0–5			0.49,
	Wethers	Saxon Merino				Linear	7,10	0.70
	Wethers	Saxon Merino	37	weaners	0–5	Linear	9.3	0.67
	Weaners, ewes	Saxon Merino			0–5	Linear	7.0	0.52
van Burgel et al. (2011)	Gestation & Lactation	Merino, Leister X Merino	1500		0–5 (0.25, 0.5)	Linear	9.2	
Sezenler et al. (2011)			156		0–5 (0.5)	Linear		
Morel et al. (2016)	Dry, ewes	Romney cross	28	4–6	1.0–5.0 (0.5)	Linear	7.7	0.66
McHugh et al. (2019)	pregnancy	Multiple breeds & crossbreds			1.0–5.0	Linear	4.9	0.14
	Lambing	Multiple breeds & crossbreds			1.0–5.0	Linear	6.3	0.18
	Lambing	Multiple breeds & crossbreds			1.0–5.0	Linear	6.3	0.18
	Pre-weaning	Multiple breeds & crossbreds			1.0–5.0	Linear	4.8	0.21
	Weaning	Multiple breeds & crossbreds			1.0–5.0	Linear	4.7	0.23
	Post-weaning	Multiple breeds & crossbreds			1.0–5.0	Linear	6.9	0.32
	Mating	Multiple breeds & crossbreds			1.0–5.0	Linear	4.1	0.23
		Belclare	540		1.0–5.0	Linear	6.4	0.16
		Charollais	1484		1.0–5.0	Linear	8.7	0.29

	Suffolk	885	1.0–5.0	Linear	6.9	0.32
	Texel	1695	1.0–5.0	Linear	5.2	0.18
	Vendeen	140	1.0–5.0	Linear	9.8	0.38

Adapted with modifications from Kenyon et al. (2014). [a]Nature of association if listed.

2.10.4 Height at withers

There are documented relationships between BCS and height at withers in sheep (Maurya et al., 2008; Holman et al., 2012; Anusha, 2016). Maurya et al. (2008) working on 247 mature Maplura ewes reported that height at withers was lower in ewes of BCS 2.5 (59.7 cm) and increased with BCS being higher in ewes of BCS 4.0 (60.5 cm). In a different study by the same author working on 119 mature Garole x Malpura ewe crosses, they reported that height at the wither was highest in ewes with BCS 3.0 (53.7 cm) compared with the ewes with BCS 2.5 (51.0 cm) and 3.5 (52.5 cm). Body condition score has been reported to be significantly correlated with height at withers in sheep (Holman et al., 2012). Both studies have reported a linear association with moderate correlation coefficients of 0.58 and 0.44 respectively. Anusha (2016), however, reported a weak correlation between BCS and height at withers of 0.28 in Nellore brown sheep. All the three studies combined suggest that BCS is linearly related to height at withers in ewes and the strength of association appears to range from weak to moderate. It is not yet known how the association between sheep BCS and height at withers varies over time. These factors affecting wither height and its association with BCS need to be investigated over time.

Physical measurements such as liveweight and wither height are positively correlated with body condition score. If the relationship between BCS and such measurements can be predicted, then, it should be possible to use such measurements as proxies for BCS singly or in a combination. Tapping into the association between BCS and these easy-to-determine variables would, therefore, establish an indirect way of generating BCS estimates.

2.11 Perspective and proposal

Although it has been shown, both LW and BCS are related to sheep performance, studies of liveweight and BCS in sheep have reported inconsistencies associated with their measurement. Liveweight measurement error has been associated with fluctuations in gut-fill based on several factors such as feed type offered, feeding level, physiological status and season. It is, therefore, of interest to gain a greater understanding of the impacts of such factors on liveweight loss profiles associated with handling and weighing of ewes. Further, it is also of interest to generate time-dependent liveweight adjusting equations that could be incorporated into weighing systems for correction of losses due to gut-fill changes which is the focus of research Chapters 3, 4 and 5. Literature has further shown that BCS which is another indicator of sheep performance has low adoption rates among farmers. Body condition score measurement employs a hands-on procedure which can be time-consuming. It is therefore of interest to explore indirect faster means of measuring BCS. A greater understanding of the relationship between BCS and physical measurements and the possibility of using such measurements as proxies for BCS in sheep may be of advantage which is the focus of research Chapters 6, 7, 8 and 9. Furthermore, evaluation of animal condition can be a complex process requiring improved accuracy of measuring both liveweight and BCS and a greater understanding of this process will be achieved when all available information is synthesized, transformed into algorithms and integrated into weighing systems.

Foreword to Chapters 3 to 5

Chapters 3, 4 and 5 of this thesis present work on factors affecting ewe liveweight loss and methodologies to correct for weight losses during delayed weighing of ewe lambs. The methodology of Chapter 3 presents liveweight loss profiles of ewe lambs offered ryegrass- and herb-clover-based swards in order to determine if the rates of weight losses differ between the two feed types. Chapter 4, examines the effect of herbage availability and season on the rate of weight loss in ewe lambs while Chapter 5, examines the effect of herbage availability and physiological state on the rate of weight loss in mixed-aged ewes. In both Chapters 4 and 5, correction equations were developed to correct for liveweight losses and provide accurate estimates of “without delay” liveweights.

Chapter 3. The effect of herbage type on the rate of liveweight loss during fasting of ewe lambs

Chapter published as: **Semakula, J, Corner-Thomas, R, Morris, S, Blair, H, Kenyon, P. (2019)**. BRIEF COMMUNICATION: The effect of herbage type prior to fasting on the rate of liveweight loss during fasting in ewe lambs. *New Zealand Journal of Animal Science and Production* 79, 131−134.

Abstract

This experiment examined the hypothesis that herbage type, would affect the rate of liveweight loss of Romney lambs after a period of fasting. Lambs (n=80) were allocated to one of two herbage types: grass (ryegrass and white clover) and herb-clover (chicory, plantain, red clover, white clover). Lambs grazed their respective treatments for one month prior to the start of the experiment. Lambs were weighed immediately after being removed from their herbage treatment and then at one-hour intervals for eight hours. Herbage type had a significant effect ($p < 0.01$) on the rate of liveweight loss over the eight-hour fast. Lambs grazing herb-clover swards had a greater ($p < 0.05$) rate of weight loss after four hours than did lambs grazing grass (0.55 vs. 0.23 kg/h, respectively). Similarly, after eight hours, lambs grazing herb-clover lost weight more rapidly ($p < 0.05$) than did those grazing grass (0.39 vs. 0.22 kg/h, respectively). These results support the hypothesis that herbage type influenced the rate of liveweight loss during fasting and indicate that farmers need to consider the type of herbage and time off herbage in order to obtain accurate liveweight data.

3.1 Introduction

Liveweight (LW) is an indicator of the current physical state of an animal, and change in LW is a useful tool in assessing how an animal is responding to its current environment (Brown et al., 2005; Wishart et al., 2017). Liveweight provides a basis for decision making regarding sheep management, therefore, accurate determination of LW is important. New advances in technology have led to commercially available automated-weighing systems. In addition, the advent of electronic weighing scales and use of radio frequency identification (RFID) make it easier to regularly collect liveweights of individuals over time (Brown et al., 2015). However, liveweight measurements can be affected by a number of factors including: growth, nutrition, health, stress, physiological state and genotype (Kenyon et al., 2014; Brown et al., 2015).

Liveweight is a measure of total body mass and includes muscle, fat, bone, organ, body fluids and gut-fill (Wishart et al., 2017). It is relatively stable over a short period of time, but alters over longer time periods in response to environmental and physiological conditions (Coates and Penning, 2000b; Wishart et al., 2017). The contents of the rumen (fluid and feed) can account for between 10 and 23% of total liveweight in ruminants (Hughes, 1976; Moyo and Nsahlai, 2018). Liveweight fluctuations due to gut-fill in ruminants are known to be affected by time since last meal, feed and water consumption, age and size of the animal, time of day relative to sunrise, ambient temperature, and differences in grazing behaviour (Hughes, 1976; Gregorini, 2012).

A number of strategies can be used to reduce liveweight variation including removal of feed and water for fixed periods of time prior to weighing, standardizing weighing procedures, taking an average of multiple liveweights in a day or across a number of successive days, weighing at a specific time relative to sunrise, standardizing the feed offered before weighing and increasing the number of animals and repetitions of the study (Coates and Penning, 2000b; Wishart et al., 2017). Such methodologies to reduce variation are time consuming and, therefore, not generally utilised except in experimental situations.

Routine on-farm sheep handling and weighing may involve many animals and mustering from fields of varying distances from the weighing location. This can result in significant delays, where individuals are held for many hours without access to food and water prior to weighing. Delays in weighing can lead to weight loss due to a reduction in gut-fill and body fluid (Burnham et al., 2009; Wilson et al., 2015). In lambs, varying levels of weight loss have been reported within flocks while waiting to be weighed. Hughes (1976) reported losses of 0.5 to 1.2 kg (1.8 to 3.8% of initial liveweight) after six hours and 1 to 1.7 kg (3.7 to 5.3% of initial liveweight) after 12 hours. Burnham et al. (2009), Wilson (2014) and Wishart et al. (2017) reported liveweight losses of 4.2 kg (9.8% of initial liveweight), 4.8 kg (7.8% of initial liveweight) and 2.9 kg (5.6% of initial liveweight), respectively after six hours. These levels of liveweight loss are likely to interfere with a comparison of live weight particularly when small liveweight changes are being investigated. Thus, there is a need for a new approach to determine and adjust for variations in live weight among animals and specific periods of time when sheep do not have access to feed and water while waiting to be weighed. The on-going improvements in weighing equipment, software and data management may offer a solution, as they have capacity for the time stamping of individual animal weights.

To date, no study has investigated the effect of diet on the liveweight loss of sheep. The aim of this study, therefore, was to investigate the effect of feed type (ryegrass-based pasture and herb-clover mix) on the rate of liveweight loss in lambs when removed from herbage.

3.2 Materials and methods

This research investigated the effect of herbage type: to profile liveweight and liveweight loss of ewe lambs offered two diets (ryegrass-based pasture and herb-clover mix), over eight hours of fasting within a handling facility. This study was a subset of another study not related to this thesis. Brief details on that study are in Appendix II.

3.2.1 Location and climate of study area

The experiment was conducted at Massey University's Keeble farm, 5 km southeast of Palmerston North (40°24' S and 175°36' E), New Zealand from April 27/2018 to April 04/2018 (late Autumn). In New Zealand, the shortest day is June 21st. Weather data for study week is presented in Appendix I Figure 1.

3.2.2 Study animal conditions, experimental design and feed management

The lambs used in this study were part of an on-going experiment (Protocol number: MUAEC 18/10). Six-month-old ewe lambs (n = 80) were allocated to one of two herbage types: an established ryegrass (*Lolium perenne*) and white clover (*Trifolium repen*) dominated sward (grass, n = 40) or a chicory *(Cichorium intybus)*, plantain (*Plantago major*), red (*Trifolium pratense*) and white clover mix (herb-clover, n = 40). Within each herbage type, half (n = 20) of the lambs were allowed access to drinking water and the other half (n = 20) restricted (Appendix II). In the current study, all subsamples from the nested study were pooled for the herbage type level analysis. The lambs were on these herbage type diets for 30 days prior to weighing. The dry matter percentage for grass and herb-clover was 22.2% and 12.2% respectively. The pasture masses were 1272.9 kg DM/ha and 1301.2 kg DM/ha for grass and herb-clover respectively.

3.2.3 Herbage mass determination, sampling, and nutritional composition

To determine the grazing herbage dry matter (DM) mass and ensure that the herbage availability levels were maintained within the desired ranges over the study period, rising plate meter heights were recorded at least two days before weighing of the ewe lambs and on the day of weighing. Herbage masses were estimated using a rising herbage plate meter (plate diameter of 355 mm; Jenquip, Feilding, New Zealand) calculated from 200 readings (R) per herbage availability level/paddock. Sward heights were calculated using plate meter readings using the equation below.

$$\text{Sward height (cm)} = [\frac{R_2 - R_1}{200}] \qquad 3.1$$

where R_2 is the final meter reading and R_1 is the rising plate meter reading before the first measurement of the plate. Sward height data collected within each paddock were converted to herbage mass according to an equation developed by Hodgson et al (1999) as shown below.

Herbage mass (kg DM/ha) = 200 + 158 x sward height (cm) 3.2

3.2.4 Liveweight measurement

The lambs were weighed in their respective treatment groups in the same sequence (i.e., first group to be weighed was always weighed first and last group last), immediately after arriving at the weighing facility from their paddock, and thereafter, they were weighed once every hour for the following eight hours. The eight-hour fasting period was considered as it fell within the stipulated time for fasting sheep in the Sheep and Beef Cattle Code of Welfare of New Zealand (Ministry of Primary Indistries, 2018). After eight hours, they were returned to their paddocks. This generated a dataset containing 640 records of liveweights, from 80 sheep. The lambs were weighed using Tru-Test™ MP600 load bars and XR5000 weigh head (Tru-Test Group, Auckland, New Zealand). The weighing system collected liveweights at a resolution of 0.1 kg for weights between 0 and 50 kg.

3.2.5 Statistical analyses

All analyses were conducted using R program version 3.4.4 (R Core Team, 2016). During the analysis, residuals were visually explored using residual plots (i.e. for potential outliers based on Cook's distances (Dhakal, 2018), for normality using qqplots and heteroscedasticity using residual vs fitted plots. Additional tests undertaken included, Shapiro-wilk test (Shapiro and Wilk, 1965; Peat and Barton, 2008) for normality and the Breush-Pagan *test for heteroscedasticity (Breusch and Pagan, 1979).* Extreme outlier and influential values were excluded from the final analysis based on their influence (Cook's Distances for the outliers greater than 4 / (Sample size – Number of predictor variables – 1), were are considered the influential points) on the final model (Hair et al., 2006). The final model residuals met the assumption of normality, linearity and homoscedasticity. There was, however, temporal autocorrelation in the residuals.

Following the data exploration, a linear mixed-effects model with polynomial time effect was fitted using “nlme”, a package for fitting regression for linear and nonlinear models (Pinheiro et al., 2018). Herbage type was fitted as a fixed variable, fasting time (linear and quadratic) as a covariate while an individual sheep effect was fitted as a random effect. Two-way herbage type x time interactions were also fitted. An autoregressive correlation structure with was fitted, to account for temporal dependency of nearby time. Effects in the model were contrasted based on Tukey’s adjustment method using the R program extensions emmeans (Russell, 2018) and multcomp (Hothorn et al., 2008) packages. Initially the maximum likelihood method was used to fit the model, after, the final model, was generated using restricted maximum likelihood (ReML) method, and relative goodness of fit determined based on Akaikes’s information criterion (AIC) values where the model with lowest value was retained. Average herbage availability (mass) was estimated using a general linear model with herbage type fitted as fixed effect.

3.3 Results

Within those models, both linear ($p < 0.001$) and quadratic time effects were significant ($p < 0.05$). There were also significant ($p < 0.05$) two-way time x herbage type and time^2 x herbage type interactions indicating differential weight loss rates. Grass being the predominant herbage in New Zealand, was used as the reference group for all comparisons. Average liveweight loss among the two herbage types did not differ significantly ($p > 0.05$) in the entire fasting time. Overall, lambs that had previously grazed the grass-based diet had a lower rate of liveweight loss compared with those on herb-clover ($p < 0.01$). Consequently, the liveweight loss rates and, thus, the prediction equations for grass and herb-clover based diets were significantly different ($p < 0.01$) (Table 3.1, Figure 3.1).

Table 3.1 Prediction parameters with standard errors in parentheses for lamb liveweight loss (kg) for herbage types (grass and herb-clover).

	Herbage type	
Parameter	Grass	Herb-clover
Initial weight	38.9±0.73	43.0±0.83
Final weight	36.9±0.68	40.4±0.80
Intercept	0.11 (0.056)	0.07 (0.068)
Time	0.28^a (0.033)	0.58^b (0.061)
$Time^2$	-0.06^a (0.005)	-0.03^b (0.005)
Adjusted R^2	0.66	0.75

[ab]: different superscripts denote significant difference at $p < 0.05$ across row. Liveweight loss predictive equations for grass-based diet (weight loss (kg) = $0.11+0.28Time - 0.06Time^2$); for herb-clover based diet (weight loss (kg) = $0.07+0.58Time - 0.03Time^2$) respectively. Model goodness of fit: the higher adjusted R^2 and lower RMSE the better. All Tests and contrasts based on Tukey's multiple comparison methods.

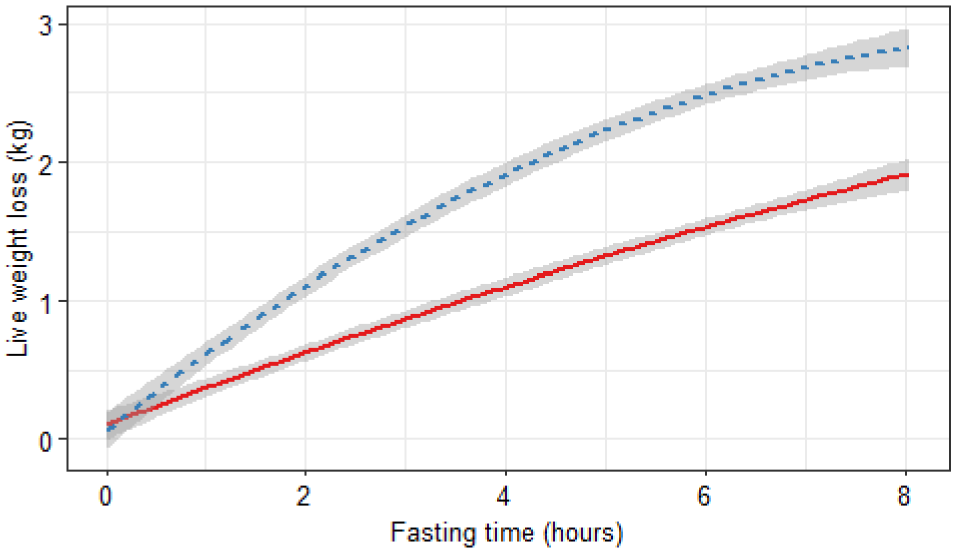


Figure 3.1 Change in liveweight (with 95% confidence interval, dotted lines) after removal from herbage, for grass (solid red line) and herb-clover (dashed blue line). Liveweight loss predictive equations for grass-based diet (weight loss (kg) = 0.11+0.28Time−0.06Time2, R^2=0.66); for herb-clover based diet (weight loss (kg) = 0.07+0.58Time–0.03Time2, R^2=0.75) respectively.

Initial liveweights for lambs on grass and herb-clover treatments were 38.9±0.93 kg and 40.2±1.03 kg, respectively. Sheep on the herb-clover treatment had a significantly greater ($p <$ 0.05) rate of liveweight loss compared with those on grass. When all data were combined, the rate of liveweight loss (0.28kg/h) was higher in the first four hours of the study compared with the later four hours (0.11 kg/h). The results further indicated that lambs fed grass lost less (0.39 kg/h., 0.22 kg/hr) weight than those on herb-clover (0.55 kg/h, 0.23 kg/h) during the first four and the entire eight hours, respectively. Results from each of the two treatment groups (grass and herb-clover), in descending order of liveweight loss, showed that lambs lost a significant amount of live weight after four (1.9±0.15 kg or 4.9% of live weight and 1.1±0.13 kg or 2.6%) and (2.8±0.16 kg or 7.3% and 1.80±0.11 4.4%) after eight hours ($p < 0.001$).

3.4 Discussion

The findings of the current study indicated that lambs lost a substantial amount of liveweight between each weighing throughout the eight-hour fasting period. The magnitude of this change is likely to influence the reliability of liveweight measures which may have implications for research and management decisions unless it can be corrected for. The lambs in the current study lost liveweight at a higher rate over the first four hours compared with the second four hours. A similar pattern of liveweight loss has been previously reported in sheep (Hughes, 1976; Wishart et al., 2017). This was previously attributed to the daily biological rhythms where the digesta from the previous day is passed from the animal in the early morning (Whiteman et al., 1954), or due to the law of diminishing returns (Wilson, 2014; Wishart et al., 2017). The liveweight losses in this study were comparable to those reported by Hughes (1976) in two-tooth sheep, Burnham et al. (2009) in hogget ewes at 10 months of age and Wishart et al. (2017) in non-pregnant dry ewes at 1.5 to 4.5 years of age, but slightly greater than those reported by Hughes (1976) in weaned and un-weaned lambs.

Lambs grazing the herb-clover diet had higher liveweight losses per unit time compared with those on the grass-based diet. This may be expected because the herb-clover mix contains a higher concentration of readily fermentable carbohydrate (soluble sugars and pectin) and lower concentrations of structural carbohydrate (i.e. cellulose and hemicellulose) than grass-based diets (Barry et al., 1999; Moyo and Nsahlai, 2018). Further, herb mixes are known to have lower Neutral Detergent Fibre (NDF) (24−49%) but correspondingly higher organic matter digestibility (68−83%) than grass-based-swards (NDF, 36−62%; OMD, 64−74%) (Golding et al., 2011; Somasiri et al., 2016) and therefore faster rumen passage (Moyo and Nsahlai, 2018). The higher hemicellulose fraction in grasses than in herbs results in higher water holding capacity and a lower digesta passage rate (Van Weyenberg et al., 2006; Moyo and Nsahlai, 2018). Hodgson et al. (1999) stated that increases in NDF concentration can restrict animal feed intake due to low rumen outflow. This suggests that the greater the hemicellulose content in the forage, the greater the amount of water it can hold. This would then result in a decrease in the fractional rate of fluid passing through the rumen and help explain the results.

To improve the reliability and comparability of liveweights it is recommended that there is standardization of feed prior to weighing of sheep and adjusting for delays. The data here provides information for farmers and scientists to correct liveweight with time off pasture. However, further work is required to validate the equations generated in the current study. In addition, further studies are needed to examine factors such as breed, age and sex of lamb, feeding levels, ambient temperature and physiological status that might interact to account for total liveweight loss.

3.5 Conclusion

For lambs fed a grass or herb-clover diet, the present study identified liveweight loss profiles during an eight-hour period when feed and drinking water were withheld. This study demonstrated that sheep lose a significant amount of liveweight over a short period and this loss rate depended on their diet type under the study conditions observed.

Chapter 4. The effect of herbage availability and season of year on the rate of liveweight loss during weighing of fasting ewe lambs

Chapter published as: **Semakula J, Corner- Thomas R.A, R, Morris S.T, Blair, H.T, Kenyon, P.R**. **(2021)**. The effect of herbage availability and season on the rate of liveweight loss during fasting in ewe lambs. *Agriculture*, 11, 150.

Abstract

Sheep (*Ovis aries*) liveweight and liveweight change can contain errors when collection procedures are not standardized, or when there are varying time delays between removal from grazing and weighing. A two-stage study was conducted to determine the effect of herbage availability and season of year on the rate of liveweight loss during fasting and to develop and validate correction equations applied to sets of delayed liveweights collected under commercial conditions. Results showed that ewe lambs offered the Low herbage availability lost up to 1.5 kg and those offered the Medium or High herbage availability lost 2.6 kg during eight hours of delayed weighing without access to feed or drinking water. The rate of liveweight loss varied by season, herbage availability and farm ($p < 0.05$). Applying correction equations on matching liveweight data collected under similar conditions, provided more accurate estimates (33–55%) of "without delay" liveweight than using the delayed liveweight. In conclusion, a short-term delay of up to eight hours prior to weighing which is commonly associated with practical handling operations significantly reduced the liveweight recorded for individual sheep. Using delayed liveweights on commercial farms and in research can have significant consequences for management practices and research results globally, therefore, liveweight data should be collected "without delay". However, when this is not feasible delayed liveweights should be corrected, and in absence of locally formulated correction equations, the ones presented in this paper could be used.

4.1 Introduction

Liveweight (LW) is an indicator of the physical state of an animal, and change in LW is a useful tool to assess how an animal is responding to its current environment (Brown et al., 2005; Wishart et al., 2017). Liveweight is a measure of total body mass and includes muscle, fat, bone, organ, body fluids and gut-fill (Wishart et al., 2017). Advances in technology have led to commercially available automated weighing systems which combine electronic scales and use of radio frequency identification (RFID). These automated systems make it more easier to regularly collect and utilize liveweight data of individuals over time (Brown et al., 2015).

Liveweight is relatively stable over a short period of time (a few minutes), but alters over longer time periods in response to environmental and physiological conditions (Coates and Penning, 2000b; Wishart et al., 2017). Liveweight measurements can be affected by a number of factors including: gut-fill (digesta and urine), growth, nutrition, health, stress, physiological state and genotype (Kenyon et al., 2014; Brown et al., 2015). The contents of the rumen (fluid and feed) can account for between 10 and 23% of total liveweight in ruminants (Hughes, 1976; Moyo and Nsahlai, 2018). Liveweight fluctuations due to gut-fill in ruminants can be affected by factors influencing feed intake such as age and size of the animal, time of day relative to sunrise, ambient temperature, and differences in grazing behaviour, and time since last meal (Hughes, 1976; Coates and Penning, 2000b; Hogan et al., 2007; Burnham et al., 2009; Gregorini, 2012; Wilson et al., 2015; Wishart et al., 2017).

In the southern hemisphere sheep production is mainly extensive in nature and pastoral based. In New Zealand, the flock sizes on average are greater than 2500 sheep (Cranston et al., 2017). Automatic weighing systems can record up 400 weights per hour without interruptions (https://www.livestock.tru-test.com), thus, requiring six to seven hours to weigh an average flock. Further, mustering and routine on-farm sheep handling in addition to weighing can increase the length of time sheep are restricted from accessing feed and water supplies while waiting to be weighed. Delays in weighing can lead to weight loss due to a reduction in gut-fill and body fluids (Hogan et al., 2007; Burnham et al., 2009; Wilson et al., 2015). In ewe lambs, varying levels of weight loss have been reported within flocks waiting to be weighed. Previously in Chapter 3, it was reported that losses of 1.8 (4.7% of initial weight) to 2.9 (6.7% initial weight) kg occur after eight hours. Hughes (1976) reported losses of 0.5 to 1.2 kg (1.8 to 3.8% of initial liveweight) after six hours and 1 to 1.7 kg (3.7 to 5.3% of initial liveweight) after 12 hours. Burnham et al. (2009) and Wishart et al. (2017) reported liveweight losses after six hours of 4.2 kg (9.8% of initial liveweight), 4.8 kg (7.8% of initial liveweight) and 2.9 kg (5.6% of initial liveweight), respectively. These levels of liveweight loss can interfere with the accuracy of comparison of liveweights, and changes in liveweight over time.

Several strategies can be used to reduce variability in liveweight including removal of feed and water for fixed periods of time prior to weighing, standardizing weighing procedures, taking multiple liveweights readings per individual per day over successive days, weighing at a specific time relative to sunrise, standardizing the feed offered prior to weighing and/or increasing the number of animals and repetitions of a study (Coates and Penning, 2000b; Wishart et al., 2017). Such methodologies to reduce variation are time consuming and, therefore, not practical for on-farm commercial use. Thus, there is a need for a new approach to determine and adjust for variations in liveweight among animals across time. The on-going improvements in weighing equipment, software and data management (Brown et al., 2015; Wilson et al., 2015) may offer a solution, as there is capacity for the time stamping of individual animal weights. Liveweight is used as a measure of an animal's productivity providing a basis for decision making regarding sheep management. Inaccurate liveweights can lead to wrong conclusions where individual animal performance or a comparison of liveweights is required. It is, thus, imperative that accurate liveweights are determined and used in sheep management.

Pre-fasting gut-fill has been found to be important in determining the rate of sheep liveweight loss during fasting (Kirton et al., 1968; Kirton et al., 1971; Thompson et al., 1987). The degree of gut-fill, retention time of particles in the gastrointestinal tract and passage rate can be affected by the quality and quantity of dry matter intake in ruminants (Alwash and Thomas, 1971; Haaland and Tyrrell, 1982; Varga and Prigge, 1982; Kaske and Groth, 1997). In Chapter 3, it was demonstrated the effect of herbage type on the rate of ewe lamb liveweight loss. It is likely that differences in the type and amount of herbage mass offered to sheep can result in variation in liveweight loss during fasting.

To date, the effect of herbage availability, season and their interaction on the liveweight loss of young sheep during fasting has not been reported. The aim of this study, therefore, was to firstly, investigate the effect of herbage availability (Low, Medium and High) and season on the rate of liveweight loss in ewe lambs when removed from herbage. Secondly, to generate and validate ewe lamb liveweight loss correcting equations. If such equations could be developed, they could then be incorporated into modern weighing systems to allow for more accurate liveweight data recording. It is hypothesized that differences in herbage availability and season would affect the rate of liveweight loss when ewe lambs were fasted.

4.2 Materials and methods

This study was conducted in two stages namely, stage one (calibration stage) which profiled the liveweight and liveweight loss of ewe lambs offered three feeding herbage availability levels (Low, Medium, and High) over two seasons (autumn and early winter), and stage two (validation stage) which evaluated liveweight loss correction equations developed from stage one, on different ewe lambs.

4.2.1 Stage one: Calibration

4.2.1.1 Location

The experimental site was at Massey University's Keeble farm, 5 km south of Palmerston North (40°24′ S and 175°36′ E), New Zealand. The experiment was conducted from 22nd March 2019 to 4th April 2019 (autumn) and repeated from 18th June 2019 to 1st July 2019 (winter). Weather data for both seasons is presented in Appendix III Figures 1a and 1b.

4.2.1.2 Study animal conditions, experimental design, and feed management

A total of 180 Romney ewe lambs were used in this study. In autumn (from 30 March to 16 April 2019), 90 ewe lambs (6–7 months of age) were selected for the study. In winter (27 May to 13 June 2019) a different group of 90 ewe lambs (8–9 months of age) were selected. The lambs were obtained from Massey University's Keeble farm and all had electronic identification ear tags (EID) and were weighed individually. The ewe lambs were randomly assigned on day one, to one of three ryegrass-based herbage availability levels; 700–900 kg DM/ha (Low herbage availability target range, n = 30), 1100–1300 (Medium, n = 30), and ≥1400 (High, n = 30) (Figure 2), ensuring that the overall groups weights were not different ($p < 0.05$). These three herbage availability levels were selected as they represented the range of potential masses these ewe lambs might be offered in normal farm practice in New Zealand. Previous studies have shown that herbage levels of 800–1000 kg DM/ha, 1200–1400 kg DM/ha have been associated with maintenance and daily liveweight gains of 120–160 g/d, respectively (Penning and Hooper, 1985; Nicol and Brookes, 2007). The herbage areas were 1.9 ha (Low), 2.1 ha (Medium), and 2.0 ha (High). Herbage availability levels were achieved by grazing the herbage using other mobs prior to allocation of study sheep. The study had the approval of Massey University ethics committee (protocol number: MUAEC 18/98).

Figure 4.1 Herbage availability (Low herbage level target range: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: >1400 kg DM/ha) offered to ewe lambs during the study time.

4.2.1.3 Liveweight measurement

Ewe lambs were weighed using Tru-Test^{TM} MP600 load bars and XR5000 weigh head (Tru-Test Group, Auckland, New Zealand) as in Chapter 3. The weighing system collected liveweights at a resolution of 0.1 kg for liveweights between 0 and 50 kg and 0.2 for weights between 50 and 100 kg. At day seven, lambs were weighed immediately after arriving at the weighing facility from their paddock (Without delay weight: within ten minutes of removal from herbage), and then again at hourly intervals (delayed weight) for the following eight hours, in their respective treatment groups which were weighed in the same group sequence. During their stay at the weighing facility, ewe lambs did not have access to feed or water. After eight hours, the ewe lambs were returned to their paddocks. This procedure occurred on two more occasions within each season, while the lambs grazed their respective herbage availability levels (autumn: day 7, day 11 and 14; winter: day 7, day 12 and 14).

4.2.1.4 Herbage sampling, mass and quality

To determine the grazing herbage dry matter (DM) mass and ensure that the herbage availability levels were maintained within the desired ranges over the study period, rising plate meter heights were recorded at least two days before weighing of the ewe lambs and on the day of weighing. Masses were estimated using a rising herbage plate meter using the procedures described in Chapter 3 (Equations 3.1, 3.2).

Herbage grab samples to represent what the lambs were consuming were collected at random for nutritional quality analysis across herbage availability levels and pooled within herbage availability level and day of collection, at days 7, 11, and 14 in autumn (n = 9) and 7, 12, and 14 in winter (n = 9). Samples were collected between 9.00 AM and 12.00 PM at each sampling time. Samples were divided into two and either freeze-dried and stored for further chemical analysis or used for dry matter determination (percentage dry matter and the ratio of live/green to dead matter).

Samples for each herbage availability level collection on each day, were mixed and a subsample of approximately 50 g fresh weight was recorded. The subsamples were then oven dried at 70 °C to a constant weight. The oven-dried herbage was then ground to pass through a 1-mm sieve and analysed for crude protein (CP), Acid detergent fibre (ADF), neutral detergent fibre (NDF), and digestible organic matter (OMD) using near-infrared reflectance spectroscopy (NIRS; Model: FOSS NIRSystems 5000, Maryland, USA) (Corson et al., 1999; Haese et al., 2020) calibrated for high water soluble carbohydrate (WSC) grasses (FeedTech, AgResearch Grasslands, Palmerston North). Additionally, a prediction of the metabolizable energy (ME) of the feed was determined using organic matter digestibility (OMD*0.16 MJ/kg) (Roughan and Holland, 1977; Dowman and Collins, 1982). The NIRS system estimates forage composition by comparing the spectral scan with a database of spectral and analytical information (predetermined from wet chemistry) to give an estimate of chemical composition (Corson et al., 1999).

The fresh herbage sample was weighed before being oven dried at 70°C for 48 hours, and then reweighed to determine its dry matter content (DM) % using the formula below.

$$\text{DM}\ \% = 100 - \left[\frac{(\text{Fresh weight} - \text{Dry weight})}{\text{Fresh weight}} \text{x } 100\right] \qquad 4.1$$

Further, a subsample (approximately 20 g) of fresh herbage was sorted into live and dead matter and then oven dried for dry matter estimation. The dry samples (live and dead) were then weighed separately to determine their dry weights. The proportion of live (green) matter to dead was calculated per herbage availability level as follows.

$$\text{Live matter}\ \% = 100 * \left(\frac{\text{Dry weight of green herbage}}{\text{Total dry weight (green+dead)}}\right) \qquad 4.2$$

4.2.1.5 Statistical analyses

All analyses were conducted using the R program version 3.4.4 (R Core Team, 2016). The data were explored and analysed using procedures in Chapter 3. During the analysis, residuals (error term) were visually explored using residual plots (i.e. for potential outliers based on cook's distances (Dhakal, 2018), for normality using qqplots and heteroscedasticity using residual vs fitted plots). Additional tests undertaken included, Shapiro-wilk test (Shapiro and Wilk, 1965; Peat and Barton, 2008) for normality and the Breusch-Pagan *test for heteroscedasticity (Breusch and Pagan, 1979).* Extreme outliers were excluded from the final analysis based on their influence (Outliers with Cook's Distances > 4/sample size, were then considered influential points) in the final model (Cook, 1977). The final model residuals met the assumption of normality, linearity and homoscedasticity. There was, temporal autocorrelation in the residuals based on visual inspection of the autocorrelation plots.

Prior to analysis, data were partitioned into two while maintaining the class balance for different herbage levels and physiological state of groups as follows; 70% of the measurements were used to train the model (training set), and the remaining 30% were used to cross-validate the model (test dataset). The ewe lamb liveweight loss training dataset was expanded using a 1000 fold cross validation resampling technique and using different splits each time, To predict ewe lamb weight loss, a mixed effects model with a first order correlation structure was fitted using R program (R Core Team, 2016) with the nlme package extension (Pinheiro et al., 2018).

Herbage availability (H) together with season (S) were fitted as fixed variables, holding time (T: first and second order polynomial) as covariate while an individual sheep effect was fitted as a random effect. Initially, all variables were fitted including their two-way (S x H, S x T, S x T^2, H x T, H x T^2) and three-way (S x H x T and S x H x T^2) interactions and then the nonsignificant ones eliminated through backward selection. The model with the least Akaike's information criterion (AIC) values (minimal model) was retained.

Herbage mass was estimated using a general linear model fitted using the generalized least squares method (GLS) in nlme package with herbage availability level, season, and sample days as fixed effect. Two-way herbage availability x season interactions nested within sample days, or three-way herbage availability x season x sample days were tested. The model with nesting structure having had the least AIC value was selected as most fitting for further analysis. All model effects were compared using the minimal model, based on Sidak's multiple-comparisons tests as in Chapter 3.

4.2.2 Stage two: Validation

4.2.2.1 Location

Data collection for the validation of the equations was conducted on two different Massey University farms (Tuapaka and Riverside), New Zealand. Tuapaka farm located 15 km north-east of Palmerston North City (40°20′ S, 175°43′ E) and Riverside farm was located 11 km north to north-west of Masterton (40°50′ S, 175°37′ E). The weather details for the validation sites are presented in Appendix III Figure 2a and 2b.

4.2.2.2 Study Animals, Experimental Design, and Feed Management

Validation was conducted using eight-month-old Romney ewe lambs (n = 90) at Tuapaka farm from the 30th July to 8th June and at Riverside farm from the 7th to 16th July (n = 90) in the winter of 2020. On day one, ewe lambs (n = 30) were randomly allocated on to one of three herbage availability levels (Low, Medium, and High as per the calibration stage). The herbage was a ryegrass and white clover-based sward mix. The herbage availability areas were 1.5 ha (Low), 3.7 ha (Medium), and 2.0 ha (High) for Tuapaka, and 3.0 ha (Low), 4.1 (Medium), and 3.1 ha (High) for Riverside.

4.2.2.3 Liveweight Measurement

Ewe lambs were placed on their respective herbage availability levels/paddocks (only one paddock per herbage availability level) for three days (days −3 to day 0) prior to start of the study. The ewes were weighed on days 4 and 6 at Tuapaka farm and days 4 and 7 at Riverside farm. Ewe lambs were weighed in their respective herbage availability groups immediately after arriving at the weighing facility from their paddock (within 10 min of removal from herbage to get the "without delay" weight), and then hourly in the same group sequence for the following six hours. During their stay at the weighing facility, ewe lambs did not have access to feed and water. After six hours, the ewe lambs were returned to their paddocks. The study had the approval of Massey University ethics committee (protocol number: MUAEC 19/53).

4.2.2.4 Herbage Sampling, Availability Determination, and Herbage Quality

Herbage availability was recorded on the first day (day one), first weighing (within 2–4 days) and last weighing day (within 5–7 days) of the study only. Herbage samples were collected on each day of weighing and analyzed for quality parameters and for dry matter percentage and proportion of live to dead matter as per the calibration stage of study.

4.2.2.5 Statistical Analyses

The validation datasets generated in stage two were collected using different groups of ewes. Two datasets, each containing 630 records of liveweights (seven weights taken in six hours including the "without delay") from 90 ewe lambs were collected at each study farm. The six-hour fasting period was considered a more practical period of delay that may occur during routine handling and weighing of a flock of sheep (Wishart et al., 2017).

To determine if the rate of liveweight loss was consistent across farms and study stages, data from the winter season in stage one, from Keeble farm (n = 1730), using up to six hours of fasting, was pooled with the two validation datasets from Tuapaka farm (n = 1078) and Riverside farm (n = 1257). A mixed effects model with a first order correlation structure was fitted using R program (R Core Team, 2016) with the nlme package extension (Pinheiro et al., 2018). All effects in the model were compared as in stage one. Herbage availability (H) and farm (F) were fitted as fixed variables, holding time (T: first and second order polynomial) as covariate, and individual ewe lamb was fitted as a random effect. All variables were fitted including their two-way (F x H, F x T, F x T^2, H x T, H x T^2) and three-way (F x H x T and F x H x T^2) interactions. Additionally, models were refitted with herbage availability and time effects nested within farm. The model without nesting structure having had the smallest AIC and BIC values (loglikelihood ratio, $p < 0.001$) was selected as most appropriate for further analysis. The nonsignificant model terms were eliminated, and the minimal model subsequently selected as in the calibration stage. Season was not considered as validation data were only collected in winter.

For prediction of average mass in each herbage availability level, a general linear model was fitted as in calibration stage with herbage availability level, farm and sample days as fixed effect. Further, two-way farm x herbage availability interactions nested within sample days, and three-way herbage availability x farm x sample days were tested. The model with nesting structure was selected as most fitting for further analysis. All model effects were compared using the minimal model.

Following the linear mixed effects regression model analysis in stage one, six separate correction equations were generated at stage one, representing each herbage availability offered (Low, Medium, and High) and season (autumn, winter). This resulted in six liveweight loss equations. The formula for computing the corrected liveweight (cW_0) is given below).

$$cW_0 = dWt + Wlt \quad 4.5$$

where, dW_t was the delayed or observed weight measurement at time (t) and Wl_t was ewe weight loss after time (t) off feed (t = 0, …, six hours) computed using the separate or consolidated weight loss equations generated in stage one.

Even though up to six hours of fasting would be the quintessential delayed time during on-farm weighing, eight hours were preferred for developing the Live weight correction equations and the subsequent without LW predictions. This is because the eight-hour-based correction equations covered more data (time points) and thus, more accurately modeled the liveweight loss trend within and after the six-hour fasting period.

Correction equations were deployed to predict the “without delay” Live weight on validation datasets collected during winter from two farms (Tuapaka and Riverside). Validation data could not be collected on ewes in autumn due to the COVID-19 lockdown imposed in New Zealand from March to June. Several metrics (Table 4.1) were used to assess the quality of models, including the coefficient of determination (R^2: multiple regression or r^2: simple regression), Lin’s concordance correlation coefficient (CCC), bias, root mean squared error (RMSE), residual prediction deviation (RPD), and the ratio of performance to interquartile distance (RPIQ) (Moriasi et al., 2007; Bellon-Maurel et al., 2010; McDowell et al., 2012; Botchkarev, 2019). The success of the predictions for individual samples was determined using the relative percent error (RPE). The best model would have the highest R^2 or r^2, CCC, RPD, and

RPIQ, and the lowest RMSE and RPE. In addition, RPD has been classified (Bellon-Maurel et al., 2010; Kodaira and Shibusawa, 2013) into three different categories, weak prediction (RPD < 1.4), reasonable (1.4 < RPD < 2.0) and excellent (RPD > 2.0). In a similar manner, RPIQ has been divided into four categories, very poor prediction (RPIQ < 1.4), fair (1.4 < RPIQ < 1.7), good (1.7 < RPIQ < 2.0), very good (2.0 < RPIQ < 2.5) and excellent (RPIQ > 2.5) (Nawar and Mouazen, 2017).

Each validation was conducted using 1000-fold cross validation (bootstrap) with three repeats. In theory as the number of times a bootstrap is conducted increases (large number of folds), the bootstrap standard deviation approximates sample standard error (Efron, 2014). Consequently, 1000 bootstraps were conducted to estimate the descriptive statistics on accuracy and error metrics (mean, standard deviation, inter-quartile range).

Table 4.1 Goodness of fit and accuracy measures of the calibration (stage one) equations applied to the validation (stage two) data sets.

Metric	Equation
Bias	$\mathrm{Bias} = \frac{1}{n}\sum_{j=1}^{n}(x_j - y_j)$
Root mean square error	$\mathrm{RMSE} = \sqrt{\frac{\sum_{j=1}^{n}(x_j - y_j)^2}{n}}$
Relative Prediction Error	$\mathrm{RPE} = \frac{\mathrm{RMSE}}{\bar{x}} * 100$
Residual prediction deviation	$\mathrm{RPD} = \frac{SD_{y_j}}{\mathrm{RMSE}}$
Ratio of performance to interquartile distance	$\mathrm{RPIQ} = \frac{(Q_3 - Q_1)}{\mathrm{RMSE}}$
Coefficient of determination	$R^2 = \frac{\sum_{j=1}^{n}(x_j y_j - \overline{xy})^2}{(\sum_{j=1}^{n} x_j^2 - \bar{x}^2)(\sum_{j=1}^{n} y_j^2 - \bar{y}^2)}$
Adjusted R^2 (Adj.R^2)	$\mathrm{Adj.}R^2 = 1 - [\frac{(1-R^2)+(n-1)}{n-k-1}]$
Lin's concordance correlation coefficient	$\mathrm{CCC} = \frac{2\rho SD_x SD_y}{SD_x^2 + SD_y^2 + (\bar{x} - \bar{y})^2}$

Where n is sample size, x_j and y_j are the actual and predicted values, respectively, and $\bar{x}$ and $\bar{y}$ are their respective means. ρ is the Pearson's correlation coefficient between the observed and predicted values. SD is standard deviation, Q1 and Q3 are the 25th and 75th quartiles respectively.

4.3 Results

4.3.1 *Calibration Stage*

4.3.1.1 *Herbage availability and Chemical Composition*

The estimated mass of available herbage (kg DM/ha), differed among herbage availability target levels ($F_{2,12} = 153.7$, $p < 0.001$) and between seasons ($F_{1,12} = 4.60$, $p < 0.05$) but not ($F_{2,12} = 0.06$, $p = 0.941$) period of study (time from day 0 to day 14) (Appendix IV Table 1a). Further, the interaction between herbage availability and season was not significant ($F_{2,12} = 0.50$, $p = 0.613$). The proportion of herbage that was considered live (green) and thus edible differed by season ($F_{1,12} = 9.4$, $p < 0.001$) and increased with herbage availability ($F_{1,12} = p < 0.05$).

The herbage chemical composition varied ($p < 0.01$) by season of year but not ($p > 0.05$) herbage availability (Appendix IV Table 1b). Dry matter, NDF, and ADF were greater ($p < 0.05$) in autumn, while, DM, CP, and ME were greater ($p < 0.01$, Appendix IV Table 1b) in winter. Within herbage availability, there were seasonal differences ($p < 0.05$) for all components. Within season, however, the herbage availability levels did not differ ($p > 0.05$) in all components.

4.3.1.2 *Effect of Herbage Availability and Season on Liveweight Loss*

Overall, the liveweight loss data were highly variable as indicated by the coefficient of variation (CV = 31–48%) (Table 4.2). The overall liveweight loss of lambs did not vary by herbage availability ($F_{2,\ 4173} = 0.53$, $p = 0.589$) or season ($F_{1,\ 4173} = 0.13$, $p = 0.722$) over the eight-hour period. However, this loss in Live weight varied linearly ($F_{1,4173} = 114.6$, $p < 0.001$) but not nonlinearly ($F_{1,4173} = 0.34$, $p = 0.558$) with fasting time. All two-way and three-way interactions were nonsignificant ($p > 0.05$) except for herbage availability x time (first order polynomial) ($F_{2,4173} = 4.35$, $p = 0.01$). After eight hours of fasting in autumn ewe lambs lost 1.54 kg (4.2% of initial weight), 1.60 kg (4.3% of initial weight), and 2.0 kg (5.3% of initial weight) for Low, Medium, and High herbage availability levels, respectively. In winter, ewe lambs lost 1.50 kg (3.2% of initial weight), 2.60 kg (4.8% of initial weight), and 2.62 kg (5.4% of initial weight) for Low, Medium, and High herbage availability levels, respectively (Figure 4.2).

The rate of liveweight loss stayed uniform (straight line) over time in autumn for each herbage availability level and for ewe lambs offered the Low herbage level during winter (Figure 4.2). However, for ewe lambs offered the Medium and High herbage levels in winter, liveweight loss was greater in the first four hours of fasting ($p < 0.05$) compared with the last four hours in winter. Although overall ewe lamb liveweight loss was comparable ($p > 0.05$) among herbage level availability levels and seasons, the rates of weight loss varied by herbage availability and season ($p < 0.01$) (Table 4.2). Generally, the rate of liveweight loss was greater ($p < 0.01$) for winter than autumn. In autumn the rate of liveweight loss was greater ($p < 0.05$) among lambs offered the high herbage level than either Medium or Low herbage level. In winter, the rate of liveweight loss was greater ($p < 0.01$) for lambs offered either High or Medium than Low herbage availability level (Table 4.2).

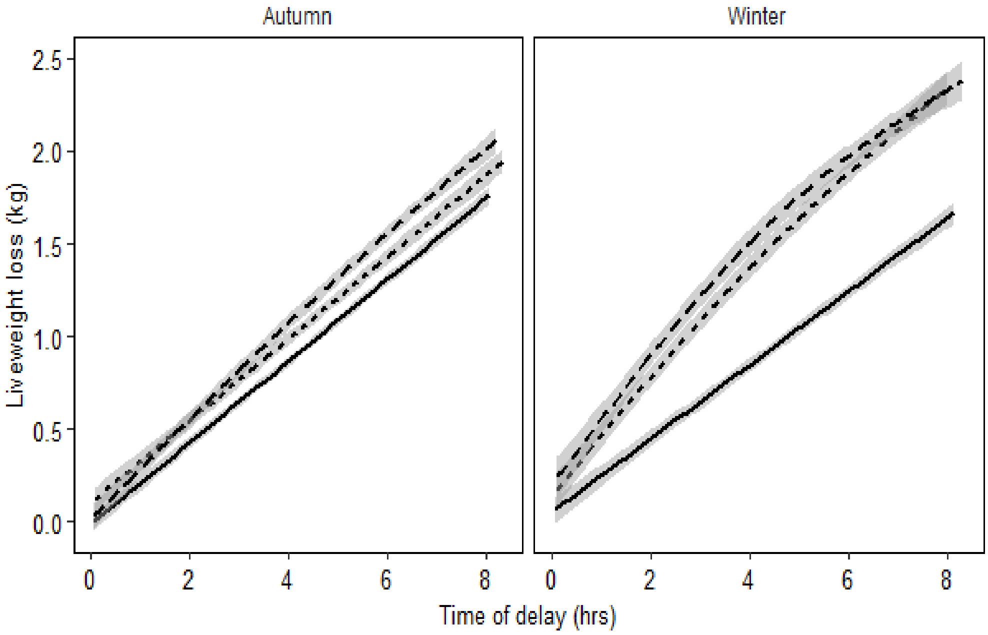


Figure 4.2 Change in Live weight (with 95% CI, grey shade) for herbage availability (Low: solid line, Medium: dashed line, High: long dashed line) in autumn and winter over eight hours of fasting.

Table 4.2 Mean initial ("without delay") and final weight and prediction parameters with standard errors in parentheses, for ewe lamb liveweight loss (kg) based on herbage availability (L: Low, M: Medium, H: High) offered to ewes by season (autumn, winter) and fasting time (1–8 hours). CV is the coefficient of variation and adjusted R^2 is a measure of goodness of fitness of the model. All models were significant at $p < 0.05$.

	Live weight (kg)		Predictor				
Herbage availability	Initial	Final	Intercept	Time	$Time^2$	Coefficient of variation (CV)	Adjusted R^2
	Autumn						
L_A	36.8±0.42	35.0±0.39	0.01(0.08)	0.20^{a}(0.02)**	ns	0.31	0.69
M_A	37.6±0.43	35.9±0.41	0.11(0.08)	0.23^{a}(0.02)**	ns	0.41	0.69
H_A	37.6±0.40	35.6±0.38	0.05(0.08)	0.27^{ab}(0.02)**	-0.020^{ab}(0.003)*	0.45	0.71
	Winter						
L_W	47.3±0.33	45.4±0.32	0.10(0.08)	0.22^{a}(0.02)**	ns	0.48	0.60
M_W	48.1±0.32	46.0±0.32	0.13(0.09)	0.35^{bc}(0.02)**	-0.012^{a}(0.002)**	0.39	0.71
H_W	48.5±0.34	46.3±0.32	0.02(0.08)	0.42^{c}(0.02)**	-0.020^{ab}(0.002)**	0.42	0.67
				Overall			
L_A+M_A+L_W	39.6(0.37)	38.8(0.36)	0.03(0.03)	0.21^{a}(0.01)**	ns	0.42	0.64
H_A	37.6(0.40)	35.6(0.38)	−0.01(0.05)	0.27^{ab}(0.02)**	$-0.0\ 2^{ab}$(0.003)*	0.45	0.71
M_W+H_W	47.5(0.23)	46.5(0.23)	0.13(0.05)	0.39^{bc}(0.02)**	−0.016(0.002)**	0.37	0.67

Initial Live weight: Live weight "without delay". Final Live weight: Live weight after eight hours of fasting. L_A+M_A+L_W and M_W+H_W are pooled combinations of herbage availability with similar regression models. Subscripts $_{A,W}$ indicate season. ^{abc} superscripts within the predictor columns (Time and Time ^{2}) per category (autumn, winter, overall), denote significant difference at $p < 0.05$. ns denotes not significant at $p > 0.05$. *, ** denote significance at $p < 0.05$ and $p < 0.01$, respectively. Model goodness of fit: the higher R^2 the better. All contrasts based on Sidak's multiple-comparisons tests.

4.3.2 Validation Stage

4.3.2.1 Herbage Mass and Chemical Composition

Herbage mass differed among herbage availability target levels and farms ($p < 0.01$) but not period between sample days ($F_{1,12} = 0.90$, $p = 0.361$) (Appendix IV 2a). Except for herbage availability x farm ($F_{2,12} = 4.48$, $p = 0.035$), all interactions between herbage availability, farm, and sample days ($p > 0.05$) were not significant. The variability (range) in herbage mass was greatest in the High availability target level. Although the aim was to maintain the herbage availability within the preset target ranges (i.e., 700–900, 1100–1300, and ≥1400 kg DM/ha), at Tuapaka farm the availability levels were slightly out of range due to unpredictable pasture growth. Consequently, Tuapaka farm had greater herbage availability levels offered to ewe lambs in both the Medium and High groups than on Riverside farm. The proportion of herbage that was considered live (green) and thus edible differed by season and increased with herbage availability ($p < 0.05$). Further, Tuapaka farm had greater live matter proportions than Riverside farm.

All herbage chemical components varied by herbage availability ($p < 0.01$) (Appendix IV 2b). There were no significant herbage availability x farm interactions ($p > 0.05$) for all chemical components. Metabolizable energy was greater at Tuapaka farm than Riverside farm ($p < 0.05$). Dry matter was lower for Medium and High herbage availability levels but was comparable ($p > 0.05$) for all herbage availability levels at Riverside farm. Crude protein and NDF increased with herbage availability while ADF decreased with increasing herbage availability across farm ($p < 0.05$).

4.3.2.2 six-hour variability in liveweight loss at calibration and validation

This section presents a comparison in the ewe liveweight loss trends during a six-hour fasting period between the calibration dataset (from Keeble farm) and two validation datasets (from Tuapaka farm and Riverside farm) (Table 4.3). The overall mean liveweight loss did not vary ($F_{2,261} = 0.54$, $p = 0.581$) between farms and among herbage availability ($F_{2,261} = 0.78$, $p = 0.460$) over the six-hour fasting period. However, the overall liveweight loss varied linearly ($F_{1,3778} = 21.43$, $p < 0.001$) but not nonlinearly ($F_{1,3778} = 0.21$, $p = 0.650$) with fasting time. All two-way time-based interactions were significant ($p < 0.001$). However, the herbage x farm interaction was not significant ($F_{1,261} = 0.45$, $p = 0.769$) and all three-way interactions were not significant ($p > 0.05$). The rate of liveweight loss varied among farms ($F_{2,3778} = 11.9$, $p < 0.001$) and among herbage availability levels ($F_{2,3778} = 46.7$, $p < 0.001$). The proportion of variance explained by each model (adjusted R^2) was greatest for Riverside farm and least for Keeble farm.

Further, the variability in data were highest at Tuapaka farm (CV = 44 – 55%) and was lowest on Riverside farm (CV = 23 – 31%).

Table 4.3 Mean initial ("without delay") and final weight and prediction parameters with standard errors in parentheses, for ewe lamb liveweight loss (kg) during a six-hour fasting period in winter, by herbage availability (Low, Medium, High) and farm (Keeble, Tuapaka, and Riverside). CV is the coefficient of variation and adjusted R^2 is a measure of goodness of fitness of the model. All models were significant at $p < 0.05$.

		Live weight (kg)		Predictor			Coefficient of Variation (CV)	Adjusted R^2
Farm	Herbage Availability	Initial	Final	Intercept	Time	$Time^2$		
* Keeble	Low	47.3 (0.33)	45.4 (0.32)	−0.28 (0.107) ^{e}	0.23 (0.033) ^{a}	0.01 (0.005) ^{e}	0.48	0.48
	Medium	48.1 (0.32)	46.0 (0.32)	0.35 (0.117) ^{cd}	0.45 (0.036) ^{c}	−0.02 (0.006) ^{cd}	0.39	0.62
	High	48.5 (0.34)	46.3 (0.32)	−0.62 (0.117) ^{bcd}	0.55 (0.036) ^{cd}	−0.03 (0.005) ^{bc}	0.42	0.58
† Tuapaka	Low	38.1 (0.26)	37.1 (0.24)	−0.62 (0.146) ^{abcd}	0.40 (0.046) ^{bc}	−0.03 (0.007) ^{bcd}	0.55	0.50
	Medium	41.5 (0.39)	39.5 (0.33)	−1.24 (0.150) ^{a}	0.80 (0.046) ^{e}	−0.06 (0.007) ^{a}	0.43	0.65
	High	42.6 (0.40)	40.3 (0.36)	−1.13 (0.153) ^{ab}	0.79 (0.047) ^{e}	−0.06 (0.007) ^{a}	0.44	0.65
† Riverside	Low	40.6 (0.45)	39.3 (0.44)	−0.13 (0.13) ^{de}	0.27 (0.041) ^{ab}	−0.01 (0.006) ^{de}	0.31	0.75
	Medium	43.8 (0.46)	41.8 (0.44)	−0.39 (0.137) ^{cd}	0.46 (0.043) ^{c}	−0.02 (0.006) ^{cd}	0.24	0.84
	High	43.9 (0.49)	41.4 (0.47)	−0.75 (0.144) ^{abc}	0.68 (0.045) ^{de}	−0.04 (0.007) ^{ab}	0.23	0.85

Initial Live weight: Live weight "without delay". Final Live weight: Live weight after eight hours of fasting. Asterisks *,† attached to farm name indicate the dataset used for the analysis (*: Calibration dataset, †: Validation dataset). ^{abcde}: different superscripts within the predictor columns (Time and $Time^2$) per herbage availability and season denote significant difference at $p < 0.05$. Subscripts $_{A,W}$ indicate season. Herbage availability (Low herbage availability target range: 700–900 kg DM/ha, Medium: 1100–1300, High: ≥1400). Model goodness of fit: the higher R^2 the better. All contrasts based on Sidak's multiple-comparisons tests.

4.3.2.3 Using Separate Correction Equations on Validation Datasets to Predict "without delay" Live weight.

The regression equations derived in the calibration phase (eight hours of fasting) were validated against two independent datasets (six hours of fasting) collected on lambs from two different farms (Tuapaka and Riverside) using the correction equations (equation 4.5). The validated results showed that the ewe lamb Live weight correction equations for all feeding levels by season developed in stage one of the present study, predicted Live weight with substantial accuracy as shown by their low RPE (0.75 – 2.93%) and high r^2 (87.9 – 99.3%) and RPIQ (3.33 – 16.8) values as compared with not using any correction method (Table 4.4, Figure 4.3).

Prediction error varied ($p < 0.05$) with time of fasting, herbage availability, and farm. The prediction error (RMSE) increased with ewe lamb liveweight loss over time (Figure 4.5). Prediction error was highest in the High herbage availability and lowest in the Low herbage availability. The prediction error was also greater ($p < 0.05$) for Tuapaka farm than Riverside farm in all herbage availability levels. Further, prediction error varied by season from which the prediction model was developed ($p < 0.05$). Live weight correcting models developed in winter were more accurate in predicting the "without delay" Live weight (i.e., directly off herbage) than those from autumn for Medium and High herbage availability levels but not for the Low herbage availability. Low herbage availability weight correcting equations had comparable accuracy or prediction error regardless of model season. Using the herbage availability and season specific correcting models to predict the "without delay" Live weight when lambs were offered the High herbage availability prior to fasting increased the prediction accuracy of the "without delay" Live weight estimates by 50.5% and 58.8% for models developed in autumn and winter, respectively, compared with using the delayed weights (not immediately off herbage). The correcting equations increased the accuracy of the "without delay" Live weight estimates in lambs offered Medium herbage availability by 48.1% and 58.8% using models developed in autumn and winter, respectively, compared with the delayed weights. The correcting equations increased the accuracy of the "without delay" Live weight estimates in lambs offered Low herbage availability by 44.1% and 41.2% using models developed in autumn and winter, respectively, compared with the delayed weights.

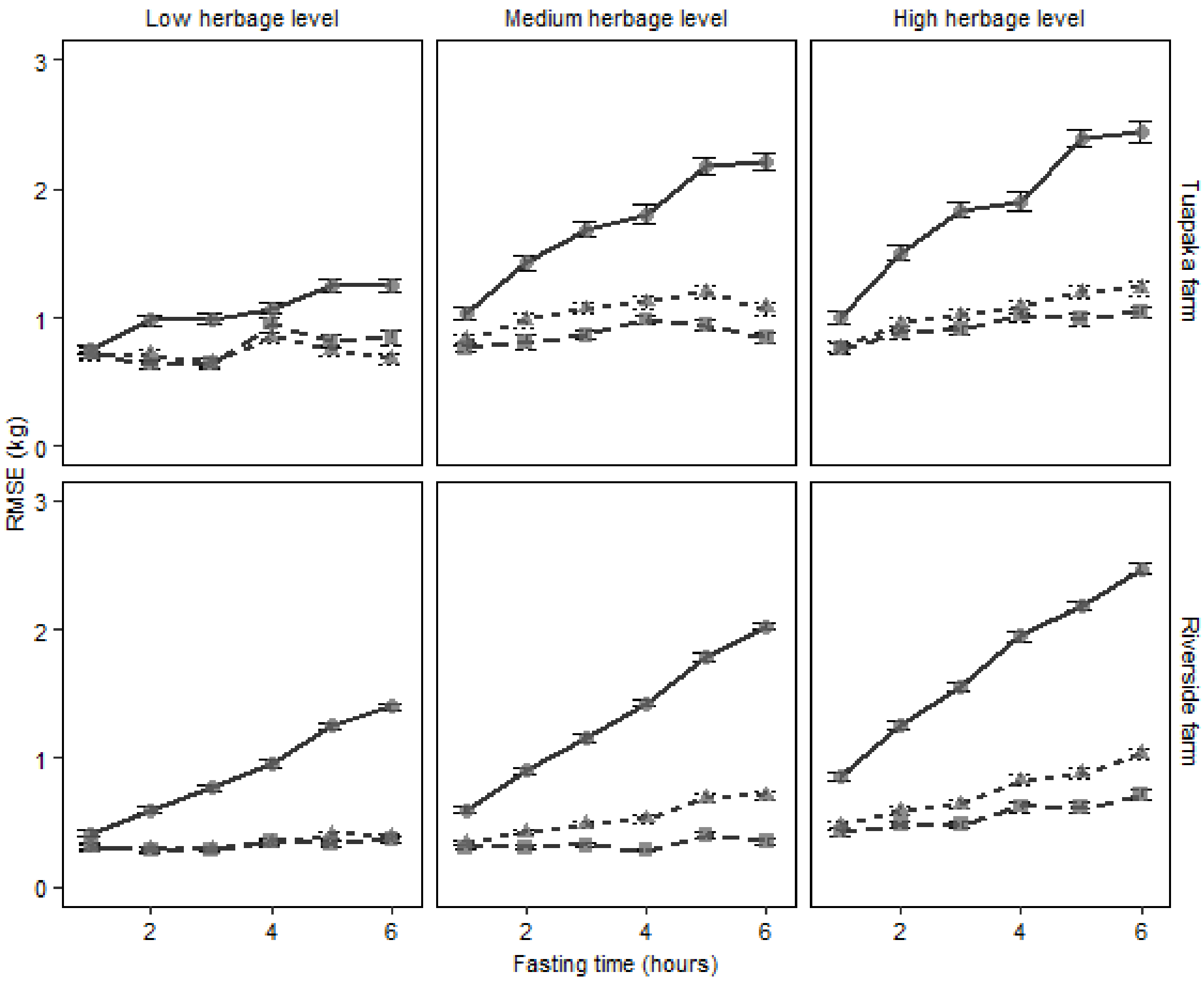


Figure 4.3 Change in root mean square error (RMSE) with associated standard deviation for the prediction of true ewe Live weight over time of fasting when individual respective correction equations (solid line with circular points: no model, dotted line with triangular points: autumn model, dashed line with square points: winter model) for each target herbage availability (Low, Medium, and High) and season generated in stage one were applied to the data collected in the winter season of 2020 to predict the “without delay” Live weight by farm (Tuapaka or Riverside). Herbage availability (Low herbage availability target range: 700–900 kg DM/ha, Medium: 1100–1300, High: ≥1400).

Table 4.4 Initial, final, and predicted live weights, measures of goodness of fit and accuracy (Bias, RMSE, RPE, RPD, RPIQ, r^2, CCC) for live weight correction models based on eight hours of fasting (from autumn and winter, 2019) applied onto the independent datasets (validation datasets) collected from Tuapaka farm and Riverside farm in Winter (2020) during a six-hour fasting period after the lambs were offered the Low, Medium, and High herbage availability.

Farm	Correction (Model)	Equation*Herbage Availability	Live weight (kg) Actual Initial	Actual Final	Predicted Final	Bias	RMSE	RPE %	RPD	RPIQ	r^2 %	CCC %
Tuapaka	None	Low	38.2 (0.26)	37.1 (0.24)		−0.75	1.06	2.78	1.92	2.45	87.9	80.8
		Medium	41.5 (0.38)	39.5 (0.33)		−1.49	1.77	4.26	1.67	2.34	80.6	75.6
		High	42.6 (0.40)	40.3 (0.36)		−1.62	1.91	4.49	1.63	2.26	81.9	73.2
	From autumn	Low	38.2 (0.26)	37.1 (0.24)	38.5 (0.24)	−0.06	0.75	1.96	2.72	3.95	87.9	92.4
		Medium	41.5 (0.38)	39.5 (0.33)	41.2 (0.33)	−0.62	1.05	2.53	2.82	3.95	92.6	94.8
		High	42.6 (0.40)	40.3 (0.36)	41.9 (0.36)	−0.58	1.06	2.49	2.93	4.08	92.6	91.8
	From winter	Low	38.2 (0.26)	37.1 (0.24)	38.6 (0.24)	0.25	0.78	2.04	2.62	3.33	87.9	91.6
		Medium	41.5 (0.38)	39.5 (0.33)	41.0 (0.33)	−0.24	0.87	2.10	3.40	4.77	92.6	94.0
		High	42.6 (0.40)	40.3 (0.36)	42.7 (0.36)	−0.35	0.94	2.21	3.31	4.60	92.6	95.5
Riverside	None	Low	40.6 (0.45)	39.3 (0.44)		−0.84	0.96	3.55	2.42	3.00	87.0	92.9
		Medium	43.8 (0.46)	41.8 (0.44)		−1.27	1.41	3.90	2.18	3.25	84.9	85.5
		High	43.9 (0.49)	41.4 (0.47)		−1.67	1.80	4.10	2.10	3.07	89.4	80.9
	From autumn	Low	40.6 (0.45)	39.3 (0.44)	40.6 (0.44)	−0.16	0.34	0.84	10.26	12.12	99.3	99.5
		Medium	43.8 (0.46)	41.8 (0.44)	43.5 (0.44)	−0.42	0.55	1.26	6.46	10.10	99.2	99.2
		High	43.9 (0.49)	41.4 (0.47)	43.0 (0.47)	−0.64	0.77	1.76	4.91	7.18	99.1	96.6
	From winter	Low	40.6 (0.45)	39.3 (0.44)	40.7 (0.44)	0.14	0.33	0.81	10.57	13.11	99.3	99.5
		Medium	43.8 (0.46)	41.8 (0.44)	43.4 (0.44)	−0.04	0.33	0.75	10.77	16.82	99.2	98.9
		High	43.9 (0.49)	41.4 (0.47)	43.8 (0.47)	−0.42	0.57	1.30	6.62	9.69	99.1	99.4

Herbage availability (Low herbage availability target range: 700–900 kg DM/ha, Medium: 1100–1300, High: ≥1400). Interpretation of measures: The best model has the highest RPD (residual prediction deviation), RPIQ (ratio of performance to interquartile distance), r^2 (coefficient of determination), CCC (Lin's concordance correlation coefficient), and the lowest root mean square error and RPE (relative prediction error). RPD (<1.4: weak, 1.4 < RPD < 2.0: reasonable, >2.0: excellent). RPIQ (<1.4: very poor, 1.4 < RPIQ < 1.7: fair, 1.7 < RPIQ < 2.0: good, 2.0 < RPIQ < 2.5: very good, >2.5: excellent). Correction equation* (None indicates delayed live weight considered).

4.3.2.4 *Using consolidated or pooled correction equations*

Measures of accuracy and changes in the standard error of prediction or the RMSE based on the combined or consolidated correction equations are given below (Table 4.5, Figures 4.4 and 4.5). Over time, the “without delay” live weight prediction accuracy was generally greater at the Riverside farm when ewe lambs were offered the Low and Medium herbage levels than on Tuapaka farm ($p < 0.05$, RMSE ± 2SD). The live weight prediction errors for High herbage level on both farms were comparable ($p > 0.05$). When applying the corresponding correction equations, the “without delay” live weight accuracies on Riverside farm were 20% higher for the High herbage level, 38% for the Medium and 37% for the Low herbage level than on Tuapaka farm (Figure 4.4). Regardless of the effect of farm or grazing location (Figure 4.5), the accuracy of the “without delay” live weight correction equations when applied to Low herbage data were consistently greatest ($p < 0.05$) for the L_A+L_W+M_A equation and H_A, compared with the M_W+H_W equation or using the delayed weights. For the Medium and High herbage levels the prediction accuracy was greatest using the M_W+H_W model and least when using the delayed weights.

The greatest improvement in accuracy when corresponding herbage availability level correction equations were used, was in the Low herbage level (55%) and lowest in the High herbage level (33%). Deploying any of the correction equations onto the delayed live weights regardless of whether they were matching with the herbage level or not, improved the “without delay” live weight estimation by 48% for the Low herbage level, 37% for the Medium and 15% for the High compared with only using the delayed live weight in place of actual weight.

Table 4.5 Measures of goodness of fit and accuracy (Bias, RMSE, MAPE, RPE, RPD, RPIQ, CCC, r^2) for overall live weight correction models (H_A: for ewe lambs offered the High herbage level in autumn, M_W+H_W: for ewe lambs offered the Medium or High herbage level in winter and $L_A+M_A+L_W$: for ewe lambs offered the Low or Medium herbage level in autumn or offered the Low herbage level in autumn) applied onto the independent datasets (validation datasets) collected on Tuapaka farm and Riverside farm in winter (2020) during a six-hour fasting period after the lambs were offered the Low, Medium and High herbage availability levels.

Farm	Model	Herbage availability	Bias	RMSE	RPE %	RPD	RPIQ	r^2 %	CCC %
Tuapaka	$L_A+M_A+L_w$	Low	0.08	0.67	1.75	4.35	5.37	96.8	97.0
		Medium	−0.67	0.98	2.36	4.49	4.54	96.5	94.3
		High	−0.79	1.07	2.51	2.91	4.12	96.7	93.3
	H_A	Low	0.17	0.69	1.81	4.21	5.2	97.2	95.2
		Medium	−0.57	0.93	2.24	4.72	4.78	97.0	96.6
		High	−0.69	1.01	2.38	3.08	4.35	97.0	95.1
	M_w+H_w	Low	0.58	0.80	2.10	3.62	4.46	77.2	79.3
		Medium	−0.16	0.79	1.90	5.57	5.63	85.7	91.7
		High	−0.29	0.85	2.01	3.64	5.15	85.8	91.5
Riverside	$L_A+M_A+L_w$	Low	−0.03	0.26	0.65	13.21	16.65	97.2	98.5
		Medium	−0.46	0.52	1.19	6.82	11.33	97.0	97.0
		High	−0.85	0.89	2.03	4.23	6.4	97.0	94.2
	H_A	Low	0.11	0.27	0.68	12.73	16.05	97.2	98.2
		Medium	−0.34	0.41	0.94	8.6	14.29	97.0	97.8
		High	−0.75	0.78	1.77	4.84	7.32	97.0	95.8
	M_w+H_w	Low	0.41	0.46	1.14	7.52	9.48	98.5	97.7
		Medium	−0.04	0.30	0.68	11.87	19.73	98.3	99.1
		High	−0.44	0.51	1.15	7.46	11.28	98.2	98.1
Overall	$L_A+M_A+L_w$	Low	−0.01	0.47	1.19	6.6	7.94	98.6	98.4
		Medium	−0.60	0.78	1.83	4.42	5.83	98.5	96.2
		High	−0.87	1.02	2.36	3.43	5.08	98.5	94.1
	H_A	Low	0.14	0.48	1.22	6.46	7.77	98.6	97.6
		Medium	−0.45	0.67	1.57	5.15	6.79	98.5	97.2
		High	−0.72	0.89	2.06	3.93	5.82	98.5	96.2
	M_w+H_w	Low	0.50	0.63	1.60	4.92	5.92	98.6	95.9
		Medium	−0.1	0.45	1.05	7.67	10.11	98.5	98.3
		High	−0.36	0.68	1.57	5.15	7.62	98.5	97.5

Herbage availability (Low herbage target range: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥ 1400 kg DM/ha). Interpretation of measures: The best model has the highest RPD (Residual prediction deviation), RPIQ (Ration of performance to interquartile distance), r^2 (Coefficient of determination), CCC (Lin's concordance correlation coefficient), and the lowest Root mean square error and RPE (Relative prediction error). Ranges for values: r^2 (0: indicates that the model explains none of the variability of the response data around its mean, 1.0 indicates that the model explains all the variability). RPD (< 1.4: weak, 1.4 < RPD < 2.0: reasonable, > 2.0: excellent). RPIQ (< 1.4: very poor, 1.4 < RPIQ < 1.7: fair, 1.7 < RPIQ < 2.0: good, 2.0 < RPIQ < 2.5: very good, > 2.5: excellent). $_{A,W}$ subscripts indicate the autumn and winter seasons respectively.

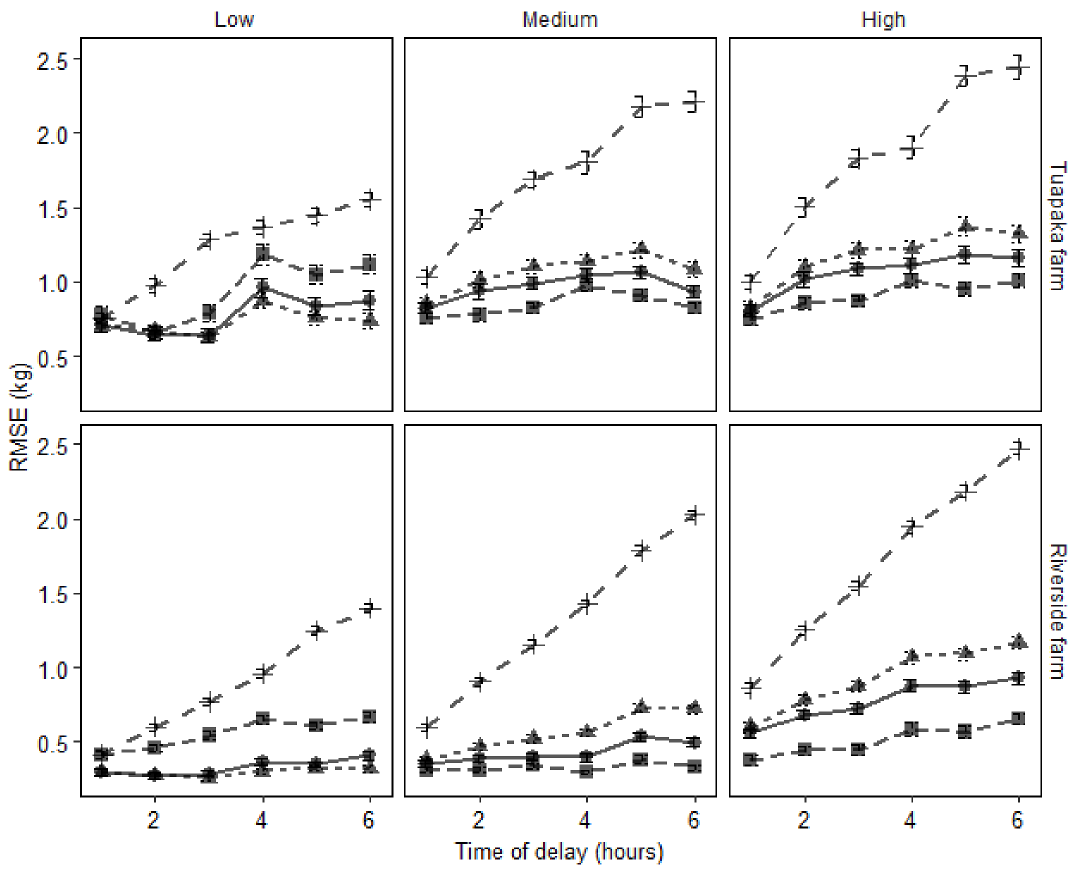


Figure 4.4 Change in root mean square error (RMSE) with associated standard deviation for the prediction of true ewe live weight over time of fasting when the consolidated correction equations (long dashed line with plus sign points: No model, dotted line with triangular points: Combined model L_A+ M_A+L_W, solid line with circular points: Model for H_A and dashed line with square points: Combined model for H_A+H_W) representative of treatments with similar rates developed in stage one were applied on data collected in winter to predict true live weight by herbage availability level (Low herbage target range: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥ 1400 kg DM/ha) and Farm (Tuapaka, Riverside). Models: H_A: for ewe lambs offered the High herbage level in Autumn, M_W+H_W: for ewe lambs offered the Medium or High herbage level in Winter and L_A+M_A+L_W: for ewe lambs offered the Low or Medium herbage level in autumn or offered the Low herbage level in Autumn. ^{A,W} subscripts indicate the autumn and winter seasons respectively.

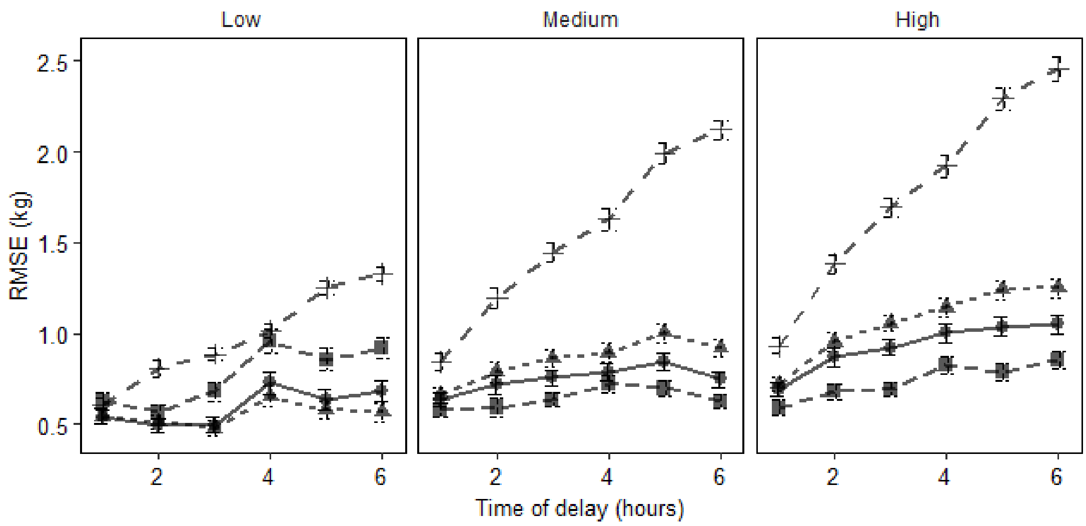


Figure 4.5 Change in root mean square error (RMSE) with associated standard deviation for the prediction of true ewe live weight over time of fasting when the consolidated correction equations (dot-dashed line with plus signed points: No model, dotted line with triangular points: Model for L_A+M_A+L_W, solid line with circular points: Model for H_A, and long dashed line with square points: Model for M_W+H_W) representative of treatments with similar rates developed in the calibration stage were applied on data collected in winter (to predict true live weight by herbage availability (Low herbage target range: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥ 1400 kg DM/ha). H_A: for ewe lambs offered the High herbage level in Autumn, M_W+H_W: for ewe lambs offered the Medium or High herbage level in Winter and L_A+M_A+L_W: for ewe lambs offered the Low or Medium herbage level in autumn or offered the Low herbage level in autumn. Subscripts $_{A,\ w}$ indicates autumn and winter respectively.

4.4 Discussion

The current study was conducted in two stages aimed; (i) to determine the effect of herbage availability and season on the rate of liveweight loss of ewe lambs during fasting and (ii) to determine if "without delay" live weight of ewe lambs could be accurately predicted from delayed live weights.

4.4.1 Calibration stage

The findings indicated that ewe lambs lost a significant amount of live weight (autumn: between 4.2 to 5.3% of initial weight, winter: 3.2 to 5.4% of initial weight) between each weighing throughout the fasting period. The magnitude of this change is likely to influence the reliability of live weight measures which may have implications for management decisions on-farm and for research unless it can be corrected for. The findings support the previous study which profiled liveweight losses of Romney ewe lambs offered grass or herb-clover-based swards in Chapter 3.

The current study indicated that the rate of liveweight loss was affected by both herbage availability and season of year, suggesting that different equations may be required to accurately correct for live weight if animals are off pasture for periods of greater than 60 min. The variation in ewe lamb liveweight loss rate by herbage availability was likely due to differences in gut-fill volume and differences in the chemical composition of the pasture (Crampton and Jackson, 1944). The DM content of the herbage was consistently lower in the High and Medium herbage availability levels than Low herbage availability level. It is, therefore, possible that the ewe lambs were consuming more water from the Medium and High herbage availability levels than the Low herbage, with this excess water being excreted faster through urine, than would herbage via fecal defecation. The seasonal differences in the chemical composition of the feeds may also have been responsible for the differential lamb liveweight loss. The lower proportions of CP and ME, but with correspondingly higher fiber (DM, NDF, and ADF) may have been responsible for the lower rate of lamb liveweight loss in autumn compared with winter. Greater structural carbohydrate and higher levels of fiber are known to increase water holding capacity in the sheep gut and thus reduce the rate of ruminal flow (Moyo and Nsahlai, 2018). In drier seasons, the proportion of fermentable carbohydrates and pectin content decrease while the structural carbohydrates (NDF and ADF) increase, however, in wetter seasons the reverse is true (Bernabucci et al., 1999; Litherland et al., 2002; Warly et al., 2004; Särkijärvi et al., 2012; Mir and Ahmed, 2017; Ekanayake et al., 2019). Therefore, it is not surprising that autumn (dry season) herbage had the highest DM and fiber and thus, the lowest rate of lamb liveweight loss compared with winter. The seasonal differences in liveweight loss could also be attributed to the higher ambient temperature experienced in autumn compared with winter during the study period. Exposure to colder temperatures can increase reticulo-rumen motility, increase the passage rate of gut particles, and reduce the gut-fill retention time (Kennedy et al., 1976; Kennedy, 1985).

4.4.2 Validation Stage

The significant polynomial regression obtained between liveweight loss and time off feed, and the subsequent linear association between delayed and “without delay” live weight, suggests there is a relationship between weight loss and “without delay” live weight. This is predicated on the hypothesis that the amount of weight lost per unit time varies depending on herbage availability and season. It was observed that the weight prediction equations tended to be more linear rather than curvilinear when animals were offered low herbage availability, or high DM%, but were curvilinear when herbage availability or when time off herbage was increased.

A comparison of liveweight loss trends using calibration and validation datasets demonstrated significant differences in overall liveweight loss and liveweight loss rates between farms. A significant farm x time interaction indicates differences in overall liveweight loss rates among farms. Further, the results indicated a greater CV % associated with this liveweight loss, which was highest at Tuapaka farm and lowest at Riverside farm. The herbage availability target ranges varied in availability levels and dry matter content which might explain the differential weight losses on different farms. Additionally, at both Keeble farm and Tuapaka farm, live weights were recorded manually by the operator whereas at Riverside farm, weights were automatically recorded. Comparison weighing was done using two 20 kg loads at the start of each weighing. However, it is possible that some error was introduced while the operator forgot to readjust the scale reading to zero each time a "shy" ewe rapidly and violently rammed into the crate gates shifting the position of the crate. An automated weighing system regularly readjusts the scale to zero, thereby reducing the error introduced due to shifts in the position of the crate.

Ideally, weighing without any delay (immediately off pasture) should provide ewe lamb live weight measurements with the least error. However, if this is not achievable, the validating process has demonstrated that correction equations can be used to supply corrected live weights (cW_0) that are more accurate estimates of the "without delay" live weight (aW_0) than a delayed live weight (dW_t). This provides a major step towards achieving improved (precise) live weight measurement in sheep production.

All correction equations were based on the eight-hour fasting period. This provided more data and, therefore, explained more variability in ewe liveweight loss. The precision of the correction equations was significantly impacted by herbage availability, season, the period of delay in recording, animal weight, and farm. This is in partial agreement with Wishart et al. (2017) who reported a significant impact of grazing location on the precision of live weight correction equations, but not time of delay before the weighing of ewes. Their study showed that the precision of the correction equations was affected by the factors associated with fluctuations in gut-fill (Coates and Penning, 2000b; Wishart et al., 2017).

The correction equations were substantially more stable when predicting "without delay" liveweight in ewe lambs offered the Low herbage availability than the Medium or High herbage availability. The consistently stable precision associated with the Low herbage availability was likely due to the higher DM% which might have caused greater water retention in the gut than for ewe lambs feeding on a lower DM% (Medium and High). In addition, the lower quantity (kg DM/ha) of herbage within the Low availability could have restricted the gut-fill thereby eliciting a response to reduce ruminal emptying. Lambs offered the High herbage availability had access to wider herbage availability ranges (1400–2200 kg DM/ha) than those offered the Low availability (700–900 kg DM/ha) which likely explains their greater error rates. However, it has been previously reported with mature ewes that intakes do not increase above a herbage availability of approximately 1400 kg DM (Morris and Kenyon, 2004).

Riverside farm had more accurate live weight estimates regarding the calibration dataset than Tuapaka farm and these differences in prediction accuracy could be explained by the variations in herbage availability levels offered to ewe lambs especially in the Medium and High herbage availability. The herbage availability estimates (Medium and High) offered to ewe lambs on the Tuapaka farm were slightly greater than those on Riverside farm. Further, the differences in prediction accuracy could also be attributed to variation in herbage dry matter percentage DM % between farms at the time of the study. This was not unexpected as Tuapaka farm is located in an area which receives more rainfall compared with Riverside farm. Overall, the results appear to suggest that increased DM % resulted in a more accurate estimate of "without delay" live weight.

In the current study, all predictions were executed on a dataset collected over one season (winter). However, all the live weight correcting equations developed from the two seasons (autumn and winter) were validated. It was not surprising that the correction equations developed for winter gave more accurate estimates than those for autumn, given the timing of the validations. However, results suggest that applying an equation from a different season to predict the "without delay" live weight from delayed live weight is a better option than using the delayed weights themselves.

The validations were conducted using a range of herbage availability levels and live weights which should cover most situations for an extensive sheep system grazing a ryegrass-based pasture. The use of simple and multiple linear regression equations based on time stamps to predict liveweight loss and to predict "without delay" live weight in sheep has been previously reported (Wishart et al., 2017). They predicted the "without delay" liveweight based on time off pasture with no reference to nutritional differences and did not provide an indication of how accurate their models were compared with not using the equation. The current study corroborates the suggestions made by Wishart et al. (2017) that the differences in quality and quantity of herbage as well as environmental factors (Hughes, 1976; Moyo and Nsahlai, 2018) which impact liveweight variation, contribute to the differences between sheep from across different farms and feeding levels.

The results of the present study demonstrated that it is possible to obtain substantially accurate estimates of "without delay" live weight of lambs offered varying availability levels of herbage prior to weighing and in different seasons of the year. It is important to correct for liveweight losses associated with handling and delayed weighing of sheep. The developed equations utilized time recorded by the weighing systems to compute the period from pasture to weighing and adjust for weight. To use these equations if incorporated into modern weighing systems, the time when ewe lambs are removed off pasture, would need to be manually entered.

4.5 Conclusion

The present study showed that ewe lambs lose a significant amount of live weight while feed and drinking water are restricted in support of findings in Chapter 3. This study demonstrated that the rate of ewe lamb liveweight loss can be predictable over a period and is dependent on herbage availability offered and season. Further, the study demonstrated that these liveweight losses can be substantially accounted for using sets of correcting equations. These equations could be incorporated into weighing systems to quickly supply farmers accurate, “without delay”, live weight measurements. Future studies should explore how to understand location-related variability in liveweight loss observed in the current study. Further, the extent to which the live weight correcting equations can be generalized to ewe lambs from different locations and breed is warranted.

Chapter 5. The effect of herbage availability on the rate of liveweight loss of fasting non-pregnant and pregnant ewes

Aspects of this Chapter have been published as: **Semakula J, Corner- Thomas R.A, R, Morris S.T, Blair, H.T, Kenyon, P.R**. The Effect of Herbage Availability, Pregnancy Stage and Rank on the Rate of Liveweight Loss during Fasting in Ewes. *Agriculture* **2021**, *11*, 543. It excludes sections on non-pregnant ewe studies.

Abstract

Sheep live weight and liveweight change are vital tools both for commercial and research farm management. However, they can be unreliable when collection procedures are not standardised, or when there are varying time delays between sheep removal from grazing and weighing. This study had two stages with different objectives: 1) a liveweight loss study, to determine the effect of herbage availability (For non-pregnant ewes: Low, Medium and High herbage levels and Low and High for pregnant ewes) on the rate of liveweight loss of ewes at different physiological states (non-pregnant and pregnant); 2) a follow-up liveweight loss study on , to develop and validate correction equations for delayed live weights by applying them to data sets collected on different farms under commercial conditions. Further, under the pregnant ewe study, ewes were evaluated at two stages of pregnancy/days of pregnancy (approximately 100 days: P100, 130 days: P130) and pregnancy-ranks (single- and twin-bearing). Results from each stage showed that the non-pregnant ewes lost up to 2.4 kg, 3.1 kg and 3.6 kg when offered the Low_1 (700 – 900 kg DM/ha), Medium (1100 – 1300 kg DM/ha) and High (≥1400 kg DM/ha) herbage levels (prior to fasting), respectively, during eight hours of delayed weighing without access to feed or drinking water. Single-bearing ewes at 100 days of pregnancy (P100) lost 3.3 kg and 5.0 kg for the Low_2 (900 –1100 kg DM/ha) and High (≥1400 kg DM/ha) herbage levels, respectively, while the twin-bearing ones lost, 3.1 kg and 4.8 kg. At 130 days of pregnancy (P130), the single-bearing ewes lost 2.8 kg and 3.5 kg for Low_2 and High, respectively, while the Twin-bearing ones lost 2.9 kg and 3.5 kg. The rate of liveweight loss varied by herbage availability, ewe physiological state and farm (p < 0.05). Applying live weight correction equations rightly (using appropriate equations on matching live weight data corrected under similar conditions) increased the accuracy of "without delay" live weight estimates in -pregnant ewes by 38% for the Low_1, 42% for the Medium and 58% for the High herbage level compared with using the delayed live weight. Within P100 ewes, the accuracy of "without delay" live weight estimates was increased by 56% and 45% for single- and twin-bearing ewes, offered the Low_2 level, respectively. The accuracy of "without delay" live weight estimates was increased by 53% and 67% for single- and twin-bearing ewes, offered the High herbage level, respectively. Among P130 ewes, the accuracy of "without delay" live weight estimates was increased by 43% and 37% for single- and twin-bearing ewes, respectively, offered the Low_2 herbage level. The accuracy of "without delay" live weight estimates was increased by 60% and 50% for single- and twin-bearing ewes, offered the High herbage level, respectively. Similarly, using a correction equation, not developed to predict "without delay" live weight using mismatching data (data collected under dissimilar conditions) gave more accurate estimates (13−60%) than using the

delayed live weight. In conclusion, a short-term delay of up to eight hours prior to weighing which is commonly associated with practical handling operations significantly reduced the live weight recorded for individual sheep. Using delayed live weights on commercial farms and in research can have consequences for management practices and research results globally, therefore, live weight data should be collected “without delay”. However, when this is not feasible delayed live weights should be corrected, and in the absence of locally formulated correction equations, the one presented in this paper could be used on farms with similar management conditions and herbage type.

5.1 Introduction

Live weight (LW) is a broadly accepted proxy for the energy status of sheep at a given time, while change in live weight is indicative of whether the animal is in either a positive energy balance (liveweight gain) or a negative energy balance (liveweight loss) (Young and Corbett, 1972; Brown et al., 2005; Wishart et al., 2017). Live weight is a measure of total body mass and includes muscle, fat, bone, organ, body fluids and gut-fill (Wishart et al., 2017). Live weight is relatively stable over shorter time periods (i.e. a few days), but alters over longer time periods in response to environmental and physiological conditions (Coates and Penning, 2000b; Wishart et al., 2017). Live weight measurements can be affected by a number of factors including: growth, nutrition, health, wool length and wetness, stress, physiological state and genotype (Kenyon et al., 2014; Brown et al., 2015). Further, the contents of the rumen (fluid and feed) can account for between 10 and 23% of total live weight in ruminants (Hughes, 1976; Moyo and Nsahlai, 2018). Liveweight fluctuations due to gut-fill (which includes the rumen and digestive tract) in ruminants are affected by factors influencing feed intake such as age and size of the animal, time of day relative to sunrise, ambient temperature, differences in grazing behaviour and time since last meal (Hughes, 1976; Coates and Penning, 2000b; Hogan et al., 2007; Burnham et al., 2009; Gregorini, 2012; Wilson et al., 2015; Wishart et al., 2017).

In countries in the southern hemisphere such as New Zealand, commercial sheep production is largely extensive in nature with flock size averages greater than 2500 sheep (Cranston et al., 2017). Commercially available automated weighing systems, combined with electronic scales and radio frequency identification (RFID), have now made it easier to regularly collect and utilize live weight data of individual animals over time (Brown et al., 2015). These weighing systems can record up to 400 weights per hour without interruptions (livestock.tru-test.com), requiring six to seven hours to weigh an average flock. Further, mustering and routine sheep handling can increase the length of time sheep are off the feed and drinking water during the weighing process. Therefore, any delays in an individual animal's weighing can lead to significant liveweight loss, due to a reduction in gut-fill and body fluids (Burnham et al., 2009; Wishart et al., 2017). In both non-pregnant and pregnant ewes, varying levels of weight loss have been reported within flocks waiting to be weighed. Burnham et al (2009) reported losses of 1.78 kg (2.7% of initial live weight) and 1.69 kg (2.6% of initial live weight) in single- and twin-bearing ewes at approximately 130 days of pregnancy after six hours, and 3.4 kg (5.3% of initial live weight) and 2.9 kg (4.5% of initial live weight) after 12 hours.

Herbage availability effects gut-fill and can influence the rate of ewe lamb liveweight loss during fasting (Chapter 4). Moreover, the physiological state of a ewe can affect intake, gut-fill and the rate of passage of fibrous food (Forbes, 1970; Rueda et al., 1990; Kaske and Groth, 1997; Moyo and Nsahlai, 2018). These levels of liveweight loss can affect the precision and accuracy of comparison of live weights and changes in live weight over time, particularly with smaller changes. Strategies aimed to reduce variability in live weight include fasting for fixed periods of time prior to weighing, standardizing weighing procedures, taking multiple live weight measurements of an individual per day or on successive days, weighing at a specific time of day relative to sunrise, standardizing the feed offered and/or increasing the number of animals and repetitions of a study (Coates and Penning, 2000b; Wishart et al., 2017). However, such methodologies to reduce variation are time consuming and, therefore, are not practical for commercial sheep farm application.

Live weight is used as a measure of an animal's productivity providing a basis for decision making regarding its management (i.e. growth rate between time points and prediction of conceptus free live weight). Inaccurate live weights can lead to incorrect conclusions where individual animal growth performance or a comparison of live weights and changes in live weight is required, for example when a ewe gains, maintains or loses conceptus free weight during pregnancy and how accurate data, coupled with known pregnancy equations can help with this. To date, it appears no studies have investigated the interaction of herbage availability and ewe physiological state (pregnancy-rank) on liveweight loss of mature ewes during fasting. Recently, it has been reported that the rate of liveweight loss in non-pregnant ewe lambs is influenced by herbage availability (Chapter 4). It was hypothesised that, a change in herbage availability offered to sheep would likely interact with its physiological status to alter the rate of liveweight loss when sheep were fasted. The aim of this study was to firstly, investigate the effects of herbage availability prior to fasting in two different reproductive stages on the rate of liveweight loss in ewes during an eight-hour period. Secondly, to generate and validate ewe liveweight loss correcting equations. If developed such equations are accurate predictors of without delay weight, they can be incorporated into modern weighing systems to allow for more accurate live weight data measurement.

5.2 Materials and methods

This research was carried out in two stages, each with different objectives. Stage one (calibration) explored the profile of live weight and liveweight loss of non-pregnant ewes offered three feeding herbage levels (Low_1: 700 – 900 kg DM/ha, Medium: 1100 – 1300, High: ≥1400), and pregnant ewes offered two feeding herbage levels (Low_2: 900 – 1100 kg DM/ha, and High: ≥1400) at 100 and 130 days of pregnancy. Stage two (validation stage) assessed the liveweight loss correcting equations established from stage one. The methodology utilized in this two-stage study was like those in Chapter 4 in ewe lambs.

5.2.1 *Calibration study*

5.2.1.1 *Location*

The study sites for the current study were like those utilised in Chapter 4 in ewe lambs. The liveweight loss profiles and correcting equations for non-pregnant ewes and ewes that were 130 days pregnant were conducted at Massey University's Keeble farm 5 km southeast of Palmerston North (40°24'S and 175°36'E), New Zealand. The study of ewes that were 100 days pregnant was conducted at Riverside farm located 11 km north to north-west of Masterton (40°50'S, 175°37'E). Weather data for the non-pregnant ewe study, Pregnant ewe study at 100 days and 130 days of pregnancy during the calibration study is presented in Appendix V as Figures 1a, 1b1 and 1b2, respectively. The study had the approval of Massey University ethics committee (protocol number: MUAEC 18/98, MUAEC 19/53). All weather data were gathered from https://cliflo.niwa.co.nz.

5.2.1.2 *Study animals, experimental design, and feed management*

5.2.1.2.1 Non-pregnant ewe study

The study of non-pregnant ewes was undertaken from 21st January to 3rd February 2020. Mixed-aged Romney ewes (3 to 5 years of age, n = 90) were allocated on day one to one of three Ryegrass-based herbage availability levels (Low$_1$ target range of: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥1400 kg DM/ha) with 30 ewes in each group (Table 5.1). A description of the herbage type used in the current study is given in Chapter 3. Previous studies have shown that herbage levels of 800 to 1000 kg DM/ha, and 1200 to 1400 kg DM/ha were associated with maintenance requirements resulting in no liveweight change and daily liveweight gains of 120 to 160 g/d, respectively (Penning and Hooper, 1985; Morris et al., 1993). The range selected represents pasture availability that ewes in New Zealand are most likely to experience during the annual production cycle (Penning and Hooper, 1985; Morris and Kenyon, 2004). The areas for each herbage availability level were 1.9 hectares, 2.1 hectares and 2.0 hectares for Low$_1$, Medium and High herbage levels, respectively.

5.2.1.2.2 Pregnant ewe study

Studies of pregnant ewes were conducted at approximately 100 days from midpoint of a 17-day breeding period (P100) and 130 days from midpoint of breeding period (P130). The P100 ewe study was conducted between 1st and 14th July 2020 on Riverside farm. The P130 ewe study was conducted between 8th and 22nd August 2019 at Keeble farm. Ewes were assigned on day one to one of the two pre-grazing herbage availability levels (Low$_2$ target range of: 900–1100 kg DM/ha, High: ≥1400 kg DM/ha). A summary of the treatment combinations, sample size and weighing days is given in Table 5.1. The areas for each herbage availability level in the P100 group were Low$_2$ (3.9 hectares) and High (4.0 hectares). The areas for each herbage availability level in the P130 group were Low$_2$ (3.9 hectares) and High (3.1 hectares).

5.2.1.3 Live weight measurement

Ewes were weighed using Tru-Test^{TM} MP600 load bars and XR5000 weigh head (Tru-Test Group, Auckland, New Zealand). The weighing system collected live weights at a resolution of 0.1 kg for live weights between 0 and 50 kg and 0.2 for weights between 50 and 100 kg. Prior to the start of the study, ewes were individually weighed to ensure they were randomly allocated to their respective groups. A summary of the calibration study conditions, weighing days and number of records is given in Table 5.1.

5.2.1.3.1 Non-pregnant ewe study

At day seven of after allocation to different herbage levels, non-pregnant ewes were weighed immediately after arriving at the weighing facility from their paddock (within five to ten minutes of removal from herbage to obtain the "without delay" weight), and then again hourly for the following eight hours. Ewes were kept in their respective herbage availability levels which the herbage availability levels were weighed in the same sequence. After eight hours, the ewes were returned to their pasture paddocks. This procedure occurred on two more occasions (at day 10 and 13), while the ewes grazed their respective herbage availability levels (Table 5.1). These weighing events generated datasets containing 2430 records (nine live weights including the "without delay" weight), from 90 ewes.

5.2.1.3.2 Pregnant ewe's study

5.2.1.3.2.1 P100 ewe study

At day six, P100 ewes were weighed immediately after arriving at the weighing facility from their paddock (within five to ten minutes of removal from herbage), and then hourly for the following eight hours in their respective herbage availability levels in the same sequence. After eight hours, they were returned to their paddocks. This procedure occurred on two more occasions (at day 14 and 16) while the ewes were grazing their respective herbage availability levels (Table 5.1).

5.2.1.3.2.2 P130 ewe study

At day seven, P130 ewes were weighed immediately after arriving at the weighing facility from their paddock (within five to seven minutes of removal from herbage), and then hourly for the following eight hours in their respective herbage availability levels in the same sequence. After eight hours, they were returned to their paddocks. This procedure occurred on two more occasions (at day 12 and 15) while the ewes were grazing their respective herbage availability levels. Datasets of 2700 records (Live weights including weight recorded immediately on arrival at the weight facility) were generated from live weights of the P100 ewes and 2700 records from the P130 ewes (Table 5.1).

5.2.1.4 Herbage sampling, mass, and quality

Herbage sampling, mass determination and quality assessment were conducted following the procedures described in Chapter 4. Herbage mass was estimated with the help of a rising plate meter (plate diameter of 355 mm; Jenquip, Fielding, New Zealand) and herbage masses were computed using the equation by Hodgson et al. (1999) as described in Chapter 3 (Equations 3.1 and 3.2).

Representative herbage grab samples were collected on the day of weight measurement (1 sample each day for each of the three herbage availability levels) at random, for nutritional quality analysis across each herbage availability level and pooled by day of collection to determine what the ewes were consuming. A total of 9 samples for the non-pregnant ewe study, 6 samples for P100 study and 6 samples for P130 ewe study were taken. Herbage samples were collected between 9.00 AM and 12.00 PM at each sampling time. Samples were divided into two and either freeze-dried and stored for further chemical analysis or used for dry matter determination (percentage dry matter and the ratio of live/green to dead matter). The samples were subjected to subsequent analyses using procedures described in Chapter 4. Sample dry matter and the proportion of live to dead were determined using Equations 4.1 and 4.2, respectively, while crude protein (CP), Neutral detergent fibre (NDF), Acid detergent fibre (ADF), Organic Matter Digestibility (OMD) and metabolizable energy (ME) were determined using the near-infrared spectroscopy (NIRS) method as described in Chapter 4.

5.2.1.5 Statistical analyses

All analyses were executed using R program version 3.4.4 (R Core Team, 2016), applying procedures described in Chapter 4. During the analysis, model residuals were explored for normality (Shapiro and Wilk, 1965; Peat and Barton, 2008) and heteroscedasticity (Breusch and Pagan, 1979). Outliers were detected based on cook's distances (Dhakal, 2018) and Cook's distances > 4/sample size were considered significant in the final model (Cook, 1977). The diagnostics above were corroborated with visual inspection of residual plots including qqplots and autocorrelation plots.

The current study data were collected on ewes of different physiological states raised in different seasons of the year, making it impossible to separate the confounding effects of season and physiological state (pregnant vs non-pregnant). Further, the experimental design was unbalanced (i.e. non-pregnant ewes were offered three levels of herbage (Low_1, Medium, High) while the pregnant ewes had two levels (Low_2, High)). Consequently, data were analysed for each of the physiological states separately.

The herbage levels were not the same across the physiological states (i.e. non-pregnant ewes were offered three levels of herbage (Low_1, Medium, High) while the pregnant ewes had two levels (Low_2, High). Therefore, for meaningful comparisons (including interactions) between and within physiological state, a separate model was fitted for each physiology state study (non-pregnant or pregnant ewe study). For each study, a mixed model (LMM) with linear and quadratic time effects and a first order auto-regressive correlation structure was fitted using R program (R Core Team, 2016) with the nlme package (Pinheiro et al., 2018). All significant model effects were reported using p-values whereas exceptional effects that were non-significant both F-and *p*-values were considered.

5.2.1.5.1 Non-pregnant ewe study

The LMM model fitted herbage availability (H) as fixed variable and fasting time (T: first and second order polynomial) as covariate and individual sheep effect as random effect. Two-way interactions between herbage availability and time (H x T, H x T^2) were also fitted. All effects were significant and were thus, retained in the final model.

5.2.1.5.1.1 Pregnant ewe study

In the pregnant ewe study, herbage availability (H), stage of pregnancy (PD) and pregnancy-rank (PR) were fitted as fixed variables, fasting time (T: first and second order polynomial) as covariate while an individual sheep effect was fitted as a random effect. Up to four-way interactions (H x PD x PR x T/T^2) were fitted. Initially the maximum likelihood method was used to build each of the models. The nonsignificant model effects were eliminated through backward selection. The model with the least Akaike's information criterion (AIC) value (minimal model) was retained. Final (minimal) models were generated using restricted maximum likelihood (REML) method. Liveweight loss equations that did not differ ($p < 0.05$) were pooled into one equation (combined).

Herbage mass was estimated using a general linear model fitted based on the generalized least squares method (GLS) in nlme package (Pinheiro et al., 2018). Herbage availability (H) was fitted as fixed effect while sample day (D) was fitted firstly, as fixed, and later as random effect. A nesting structure for herbage availability nuzzled within sample day was also investigated. The model with nesting structure having had the least AIC and BIC values was selected as most fitting for further analysis. All model effects were compared using the minimal model. Model effect means were compared based on Sidak's adjustment method (Alberts and Abdi, 2007) using the R program extensions emmeans (Russell, 2018) and multcomp (Hothorn et al., 2008) packages.

Prior to analysis, data were apportioned into two sets (70% of the measurements: training dataset for model training; 30% of the measurements: test dataset for cross-validation). Model parameters were calculated and compiled through a 1000-fold bootstrapping technique.

Table 5.1 Sample size, weighing day (D) (pregnancy day in parenthesis) and live weight records of ewes by study stage, physiological state, stage of pregnancy, pregnancy-rank, farm, and herbage availability.

Study stage	Physiological state	Stage of pregnancy	Pregnancy-rank	Farm	Herbage availability	Ewe/paddock (n)	Weighing day 1	Weighing day 2	Weighing day 2	Records
Stage 1 (Calibration)	Non-pregnant			Keeble	Low$_1$	30	D7	D10	D13	810
					Medium	30	D7	D10	D13	810
					High	30	D7	D10	D13	810
	Pregnant	P100	Single	Riverside	Low$_2$	25	D7 (107)	D10 (110)	D16 (116)	675
					High	25	D7 (107)	D10 (110)	D16 (116)	675
			Twin		Low$_2$	25	D7 (107)	D10 (110)	D16 (116)	675
					High	25	D7 (107)	D10 (110)	D16 (116)	675
		P130	Single	Keeble	Low$_2$	25	D7 (127)	D12 (132)	D15 (135)	675
					High	25	D7 (127)	D12 (132)	D15 (135)	675
			Twin		Low$_2$	25	D7 (127)	D12 (132)	D15 (135)	675
					High	25	D7 (127)	D12 (132)	D15 (135)	675
Stage 2 (Validation)	Non-pregnant			Keeble	Low$_1$	30	D3	D6		540
					Medium	30	D3	D6		540
					High	30	D3	D6		540
				Riverside	Medium	30	D4			540
	Pregnant	P100	Single	Tuapaka	Low$_2$	25	D3 (102)	D5 (107)		450
					High	25	D3 (102)	D5 (107)		450
			Twin		Low$_2$	25	D3 (102)	D5 (107)		450
					High	25	D3 (102)	D5 (107)		450
			Single	Keeble	Low$_2$	25	D3 (98)	D5 (100)		450
					High	25	D3 (98)	D5 (100)		450
			Twin		Low$_2$	25	D3 (98)	D5 (100)		450
					High	25	D3 (98)	D5 (100)		450
		P130	Single	Keeble	Low$_2$	25	D3 (127)	D5 (132)		450
					High	25	D3 (127)	D5 (132)		450
			Twin		Low$_2$	25	D3 (127)	D5 (132)		450
					High	25	D3 (127)	D8 (132)		450

Single	Tuapaka	Low$_2$	25	D3 (127)	D8 (132)	450
		High	25	D3 (127)	D8 (132)	450
Twin		Low$_2$	25	D3 (127)	D8 (132)	450
		High	25	D3 (127)	D8 (132)	450
Single	Riverside	Low$_2$	25	D6 (129)	D8 (131)	450
		High	25	D6 (129)	D8 (131)	450
Twin		Low$_2$	25	D6 (129)	D8 (131)	450
		High	25	D6 (129)	D8 (131)	450

Stage of study (Stage 1: calibration phase for profiling ewe liveweight loss based on eight hours of fasting as well as development of the correction equations, Stage 2: validation phase for evaluation of live weight correction equations when applied on other data collected different ewes based on six hours of fasting). Stage of pregnancy (P100: 100 days of pregnancy from the midpoint of a 17-day breeding period, P130: 130 days). Herbage availability (for non-pregnant ewes, Low$_1$ target range herbage availability of: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥1400 kg DM/ha; for pregnant ewes, Low$_2$ target range herbage availability of: 900 − 1100 kg DM/ha, High: ≥ 1400 kg DM/ha;). Pregnancy-rank (Single: single-bearing, Twin: twin-bearing). Weighing (values outside parenthesis denote the day of trial/study i.e. from day of animal allocation to the herbage levels while those in parenthesis indicate the average number of days of pregnancy from the midpoint of a 17-day breeding period). Ewes were weighed within 10 minutes after removal from pasture and water on Keeble farm and Tuapaka farm, and within 15 minutes for Riverside farm). Ewes were weighed in their respective herbage availability levels immediately after arriving at the weighing facility from their paddock, and then hourly in the same herbage availability level sequence for the following eight (calibration) or six (validation) hours

5.2.2 Validation study

5.2.2.1 Location

The study was approved by Massey University ethics committee (protocol number: MUAEC 19/53). The validation phase was carried out at Massey University's Keeble farm, Tuapaka farm and Riverside farm. The locations of both Keeble farm and Riverside farm were used in the calibration stage, while Tuapaka farm was located 15 km north east of Palmerston North city (40°20'S, 175°43'E). The current study stage utilized sites like those in Chapter 4 validation stage. The weather data for the three different studies and different farms during the validation study is presented in Appendix V (Figures 2a, 2b1, 2c). All weather data were gathered from https://cliflo.niwa.co.nz.

5.2.2.2 Study animals, experimental design, and feed management

The liveweight loss equations generated at calibration, were used to develop the "without delay" live weight correcting equations for the respective physiological states. The validation was conducted using both non-pregnant (n = 90) and pregnant (n = 100) mixed-aged ewes. The sample size used in the current study was based on a 0.91 power (effect size: 0.48 and non-sphericity: 0.70), it is therefore, it was sufficient to detect any effects and or differences between treatment effects.

5.2.2.2.1 Non-pregnant ewe study

Two validation studies were conducted using mixed-aged (3-5 years old) non-pregnant ewes not previously used in study one (calibration) (Table 5.1). The first study was conducted from 3rd to 10th February 2020 at Keeble farm. Ninety ewes (n = 90) were initially individually weighed. These were allocated (day one) to one of three herbage levels (Low_1 target herbage range: 700–900 kg DM/ha; Medium: 1100–1300 kg DM/ha; High: ≥1400 kg DM/ha) with 30 ewes in each group. The resulting ewe herbage level groups had similar average weight. The herbage availability areas were 1.9 hectares, 2.1 hectares and 2.0 hectares for Low_1, Medium and High herbage levels respectively. A second non-pregnant ewe validation was conducted at Riverside farm from 10th to 14th March 2020 using 30 ewes offered the Medium herbage level of area 14.9 hectares. It was not possible to obtain all three herbage levels due to a scarcity of pasture.

5.2.2.2.2 Pregnant ewe study

5.2.2.2.2.1 P100 ewe study

Two validation studies were conducted using mixed-aged ewes at 100 days of pregnancy not previously used in study one (calibration) (Table 5.1). In the first study, ewes were studied from 7th to 14th July 2020 at Keeble farm (n = 100) and in the second study from 23rd to 29th July at Tuapaka farm (n = 100). During each study a hundred ewes (n = 100) were initially individually weighed and then allocated (day one) to one of two herbage availability levels (Low$_2$ (L$_2$) target range of 900–1100 kg DM/ha, High (H): ≥1400 kg DM/ha) with 50 ewes (25=single and 25=twin-bearing) in each group, such that the overall group live weight mean did not differ. The herbage availability areas were 3.7 hectares and 2.0 hectares for Low$_2$ and High herbage levels respectively at Tuapaka farm and 3.9 hectares and 3.0 hectares for Low$_2$ and High herbage levels respectively at Riverside farm.

5.2.2.2.2.2 P130 ewe study

Three validation studies were conducted using mixed-aged ewes at 130 days of pregnancy not previously used in study one (Table 5.1). The first study was conducted from 24th to 31st July 2020 on Riverside farm (n = 100), the second study from 1st to 10th August 2020 at Keeble farm (n = 100) and the third study from 14th to 24th August 2020 on Tuapaka farm (n = 100). During each study a hundred ewes (n = 100) were initially individually weighed and then allocated (day one) to one of two herbage availability levels (Low$_2$ (L$_2$) target range of 900–1100 kg DM/ha, High (H): ≥1400 kg DM/ha) with 50 ewes (25 = single and 25 =t win bearing) in each group. The herbage availability areas were 2.1 hectares and 2.7 hectares at Keeble farm, 2.0 hectares and 3.7 hectares at Tuapaka farm, and 3.9 hectares and 3.0 hectares at Riverside farm for Low$_2$ and High herbage levels, respectively. The ewes had access to herbage and water *ad lib* up to the time they were picked from the paddock for the initial weighing.

In all pregnant ewe validation studies, ewes had been bred over a 17-day period and half were carrying single and the other half (n = 50) twin pregnancies. The ewes were placed on their respective herbage availability/paddocks (only one paddock per herbage availability level) for three days (days −3 to day 0) prior to start of the study.

5.2.2.3 Live weight measurement

Ewes were weighed as in stage one during six hours of fasting. The six-hour fasting period was considered a more practical period of delay that may occur during routine handling and weighing of a flock of sheep (Wishart et al., 2017). Without delay live weight was defined as weight taken immediately on arrival at the weighing facility from paddock. A summary of the validation study conditions, weighing days and number of records is given in Table 5.1. The ewes were weighed on two occasions on Keeble farm and at one occasion for Riverside farm. Ewes at approximately 100 days (P100) of pregnancy were weighed on two occasions two days apart on Keeble farm and Tuapaka farm. At approximately day 130 (P130) of pregnancy, ewes were weighed on two occasions two days apart on Keeble farm, five days apart for Tuapaka farm and three days apart for Riverside farm. After each day's weighing, the ewes were returned to their paddocks.

5.2.2.4 Herbage sampling, mass, and quality

Herbage mass determination and target range monitoring over the study period, was conducted as in calibration with sward height measurements recorded twice (on each day ewes were weighed) using a rising herbage plate meter. Representative herbage grab samples were also collected and analysed for quality parameters as per stage one. A total of 6 samples for non-pregnant ewes, 4 samples for P100 and 4 samples for P130 ewes were collected.

5.2.2.5 Statistical analysis

The current study data were collected on ewes of different physiological states raised on different farms, in different seasons, making it impossible to account for the confounders of ewe liveweight loss. Therefore, data were analysed for each of the physiological states as in stage one.

5.2.2.5.1 Non-pregnant ewe study

In the non-pregnant ewes, two datasets, each containing 630 (from ewes offered the Low, Medium and High herbage levels) live weights (7 weights taken in six hours including the "without delay") from 90 ewes were collected at Keeble farm and one dataset containing 210 live weights (from ewes offered the Medium herbage level) from Riverside farm from 30 ewes. A mixed effects model including the fixed effects of Herbage availability (H) (only for Keeble farm), and fasting time (first (T) and second order (T^2) polynomial) was fitted to the data with measurement days as replicates. Two-way (H x T, H x T^2) interactions were also tested for Keeble farm. To estimate the herbage availability and quality in the non-pregnant ewe study, herbage levels and chemical composition parameters were averaged to obtain the overall Medium herbage availability. In the pregnant ewe study, a general linear model was used after adjusting for herbage availability as fixed effects while day of collection was considered a replicate to estimate both availability and chemical composition.

5.2.2.5.2 Pregnant-ewe study

In the pregnant ewes, two datasets, each containing 700 live weights (100 ewes, 7 weights taken in six hours including the "without delay"), from 100 ewes (in both 100 and 130 days of pregnancy) were collected at each farm (Keeble and Tuapaka). Data from stage one, from Riverside farm for P100 (n = 1730), using up to six hours of fasting and from Keeble farm for P130 (n = 1730), was pooled with two P100 (Keeble farm: n = 1260; Tuapaka farm: n = 1302) and three P130 (Keeble: n = 1200; Tuapaka farm: n = 1078; Riverside farm: n = 1257) validation datasets. A mixed effects model with a first order correlation structure was fitted to the resulting dataset. Study stage (K), Herbage availability (H), physiological state (P), pregnancy-rank (R), stage of pregnancy (S) and farm (F) were treated as fixed variables, fasting time (first (T) and second order fasting time polynomial (T^2)) as covariate while an individual ewe effect was fitted as a random effects. Initially a mixed effects model was fitted including all main effects and up to five-way interactions (K x H x R x S x T and K x H x R x S x T^2) nested within farm (because the data were not balanced). Later, a separate model was fitted for each of the stages of pregnancy to include the fixed effect of farm. Each of the two models were fitted with up to five-way interactions (K x H x F x R x T; K x H x F x R x T^2). Minimal models were selected using the least AIC and BIC.

Following the linear mixed effects model analysis in the calibration stage, eleven separate correction equations were generated during stage one, representing each herbage availability offered (Low_1, Medium and High for non-pregnant ewes; Low_2 and High for pregnant ewes) and physiological state (Non-pregnant and pregnant) and days of pregnancy (P100, P130) (Table 5.2). To predict the ewe live weight immediately after leaving the paddock ("without delay" live weight), correction equations were developed for each of the eleven liveweight loss equations. The "without delay" live weight (cW_0) was calculated as in Chapter 4.

Live weight correction equations were deployed to predict the "without delay" live weight on validation datasets collected from the farms. These correction equations were applied on corresponding (i.e. collected under similar conditions: Separate) or on a non-matching (Mistaken) dataset, and the "without delay" live weight predictions compared with the delayed weights (where no prediction equations were used: None). Additionally, the equations that were not different (p < 0.05, Table 5.2), were consolidated to give the combined equations (Combined) which were also applied to marching datasets. It was not possible to collect data from all farms due to herbage scarcity (non-pregnant ewes) and synchronized breeding. Several metrics (Table 5.2) were computed to assess the quality of the live weight correcting equations as in Chapter 4.

5.3 Results

5.3.1 Calibration stage

5.3.1.1 Herbage mass and proportion of live/green matter

Overall, the recorded herbage masses were within the target ranges for all herbage availability levels and physiological states (pregnancy status), stages of pregnancy and pregnancy-ranks except for the Low_1 herbage level offered to non-pregnant ewes at Keeble farm (Appendix VI Table 1a). Average herbage mass also varied between herbage availability levels (p

< 0.05) for both non-pregnant ($p < 0.001$) and pregnant ewes ($p < 0.001$). In the pregnant ewe study, herbage availability did not vary between stage of pregnancy (P100 vs P130) ($F_{1,21} = 1.41$, $p = 0.248$). In the non-pregnant ewe study, the Low_1 herbage levels were slightly outside (above) the target range. In the pregnant ewe study, herbage levels were within the target ranges with the High herbage level having consistently greater ($p < 0.05$) masses than the Low_2 level, irrespective of stage of pregnancy.

Overall, the proportion of herbage that was considered live (green) and thus edible increased with herbage availability ($p < 0.05$) in non-pregnant ewes but was comparable for all herbage levels in pregnant ewes (Appendix VI Table 1a). The herbage levels offered to pregnant ewes had greater proportions ($p < 0.01$) of live or green herbage than that offered to non-pregnant ewes. Within the non-pregnant ewes, the proportions of live/green matter were as low as 57% for the Low_1 herbage level and as high as 65% for the High herbage level.

5.3.1.2 *Herbage chemical composition*

In the non-pregnant ewe study, all herbage chemical composition parameters did not vary ($p > 0.05$) with herbage availability level except for % DM ($p < 0.05$) (Appendix VI 1b). The herbage levels offered to non-pregnant ewes had greater Dry matter, NDF and ADF, but were correspondingly lower in CP and ME compared with that offered to their pregnant counterparts.

In the pregnant ewe study, all chemical composition parameters varied ($p < 0.05$) with herbage availability except ($p > 0.05$) for DM and CP. Among pregnant ewes, dry matter was greater for the Low (Low_1 and Low_2) levels than High herbage level ($p < 0.05$). Further, within pregnant ewes, herbage chemical composition did not differ among stage of pregnancy studies.

5.3.1.3 *Effect of herbage availability and physiological state on the overall and rate of liveweight loss*

Overall, liveweight loss varied over fasting time (6-8 hours) ($p < 0.01$) , but not pregnancy-rank of a ewe ($p > 0.05$). The ewes lost liveweight over the fasting period in all studies (Table 5.2, Figures 5.1 and 5.2). The overall liveweight losses over the eight-hour fasting period for non-pregnant ewes were 2.4 kg (3.8 % of the initial live weight), 3.1 kg (4.7%) and 3.6 kg (5.3%) for Low_1, Medium and High herbage levels, respectively. In the pregnant ewes, single-bearing ewes at P100 lost 3.3 kg (5.1% of the initial live weight) and 5.0 kg (7.2%) for the Low_2 and High herbage levels, respectively. Twin-bearing ewes at P100 lost, 3.1 kg (4.5%) and 4.8 kg (6.5%) for Low_2 and High. Single-bearing ewes at P130 lost 2.8 kg (4.0%) and 3.5 kg (4.8%) for Low_2 and High, respectively. Twin-bearing ewes at P130 lost 2.9 kg (4.0%) and 3.5 kg (4.6%) for the Low_2 and High herbage levels, respectively. Additionally, the variability in liveweight loss data were comparable (CV ± 2SD) across physiological state (CV = 20 − 31%) and herbage levels except for the Low_1 (CV = 58%) and medium (CV = 47%) herbage level s offered to non-pregnant ewes.

The liveweight loss regression equations differed by physiological state and herbage availability resulting in different rates of loss and the loss was nonlinear ($p < 0.001$) over the eight-hour fasting period (Table 5.2). Further, non-pregnant ewes had lower liveweight loss rates than their pregnant counterparts ($p < 0.01$). The rate of ewe liveweight loss increased with herbage availability, having been greatest in ewes offered the High rather than Low (Low_1, Low_2) herbage levels ($p < 0.01$). In the pregnant ewe study, the rate of ewe liveweight loss was greater at P100 than P130 ($p < 0.01$). Both herbage availability and stage of pregnancy significantly ($p < 0.01$) interacted to influence the rate of ewe liveweight loss. Pregnancy-rank did not affect ($p > 0.05$) the rate of ewe liveweight loss, and thus, individual liveweight loss regression equations for single- and twin-bearing ewes were pooled to generate the combined or consolidated equations.

Table 5.2 Mean initial and final delayed weight and prediction parameters with standard errors in parentheses and adjusted R^2 for ewe liveweight loss (kg) based on herbage availability level s (Low, Medium, High) offered to ewes of two physiological states (non-pregnant, pregnant), stage of pregnancy (P100: 100 days of pregnancy from the midpoint of a 17-day breeding period, P130: 130 days), pregnancy-rank (S: single-bearing, T: twin-bearing) during eight hours of fasting. Adjusted R^2 is the goodness-of-fit of the model.

Physiological state	Herbage availability	Initial live weight (kg)	Final live weight (kg)	Intercept	Time	Time^2	Coefficient of variation (CV)%	Adjusted R^2
Non-pregnant	Low$_1$	64.0(0.82)	61.6(0.80)	0.10(0.153)	0.21^{a}(0.038)	0.008^{b}(0.005)	0.58	0.58
	Medium	65.3(0.72)	62.2(0.70)	0.17(0.158)	0.32^{b}(0.039)	0.005^{ab}(0.005)	0.47	0.63
	High	67.9(0.92)	64.3(0.89)	0.09(0.162)	0.54^{c}(0.040)	−0.011^{a}(0.005)	0.31	0.76
Pregnant	*P100*							
Single (S100)	Low$_2$	64.1(0.79)	60.8(0.81)	−0.683(0.089)	0.65^{ab}(0.015)	−0.030^{a}(0.002)	0.23	0.72
	High	69.7(0.85)	64.7(0.81)	−0.688(0.089)	0.78^{c}(0.015)	−0.031^{a}(0.002)	0.30	0.67
Twin (T100)	Low$_2$	69.2(0.90)	66.1(0.81)	−0.694(0.089)	0.64^{ab}(0.016)	−0.030^{a}(0.002)	0.24	0.75
	High	73.8(1.09)	69.0(1.81)	−0.768(0.090)	0.77^{c}(0.015)	−0.032^{a}(0.002)	0.31	0.67
	P130							
Single (S130)	Low$_2$	69.9(0.88)	67.1(0.87)	−0.727(0.090)	0.59^{a}(0.015)	−0.030^{a}(0.002)	0.32	0.72
	High	72.5(0.91)	69.0(0.86)	−0.706(0.089)	0.70^{b}(0.015)	−0.030^{a}(0.002)	0.31	0.72
Twin (T130)	Low$_2$	72.9(0.69)	70.0(0.65)	−0.704(0.089)	0.61^{a}(0.015)	−0.030^{a}(0.002)	0.29	0.63
	High	76.4(0.87)	72.9(0.82)	−0.649(0.090)	0.70^{b}(0.016)	−0.031^{a}(0.002)	0.28	0.77
Combined (overall)								
(ST)P100	Low$_2$	66.4(0.64)	63.2(0.62)	0.11(0.064)	0.64^{ab}(0.027)	−0.025^{a}(0.003)	0.23	0.73
(ST)P100	High	71.7(0.71)	66.8(0.69)	−0.06(0.068)	0.81^{d}(0.030)	−0.037^{a}(0.004)	0.30	0.81
(ST)P130	Low$_2$	71.4(0.57)	68.5(0.55)	0.06(0.055)	0.56^{a}(0.023)	−0.027^{a}(0.002)	0.30	0.68
(ST)P130	High	74.5(0.65)	70.9(0.62)	−0.04(0.060)	0.71^{d}(0.019)	−0.039^{ab}(0.002)	0.29	0.75

^{abcd}: different superscripts denote significant difference at $p \leq 0.05$ in a column per physiological state. Herbage availability (for non-pregnant ewes, Low_1 target range herbage of: 700–900 kg DM/ha, Medium: 1100 −1300 kg DM/ha, High: ≥ 1400 kg DM/ha; for pregnant ewes, Low_2 target range herbage of: 900 − 1100 kg DM/ha, High: ≥ 1400 kg DM/ha;). Pregnancy-rank (S: single-bearing, T: twin-bearing, ST: Combination of single- and twin-bearing ewes). Stages of pregnancy (P100: 100 days of pregnancy from the midpoint of a 17-day breeding period, P130: 130 days). The best model has the highest R2 and the lowest RMSE. The best model has the highest R^2 and the lowest RMSE.

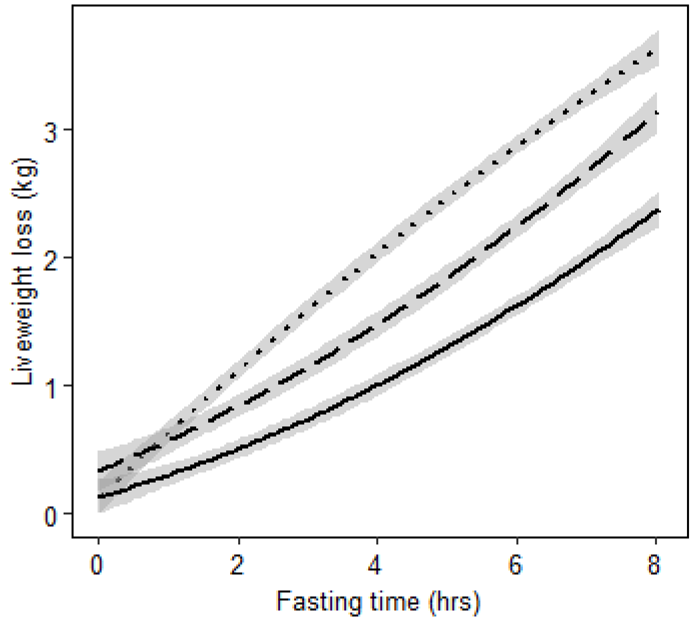


Figure 5.1 Change in live weight (with 95% CI, grey shade) of ewes offered the Low_1 target range herbage availability of (solid line), Medium (long dashed line) and High (dotted line) herbage levels during fasting of pregnant ewes. Herbage availability (Low1 target range herbage availability: 700–900kg DM/ha; Medium: 1100–1300 kg DM/ha; High: ≥1400 kg DM/ha).

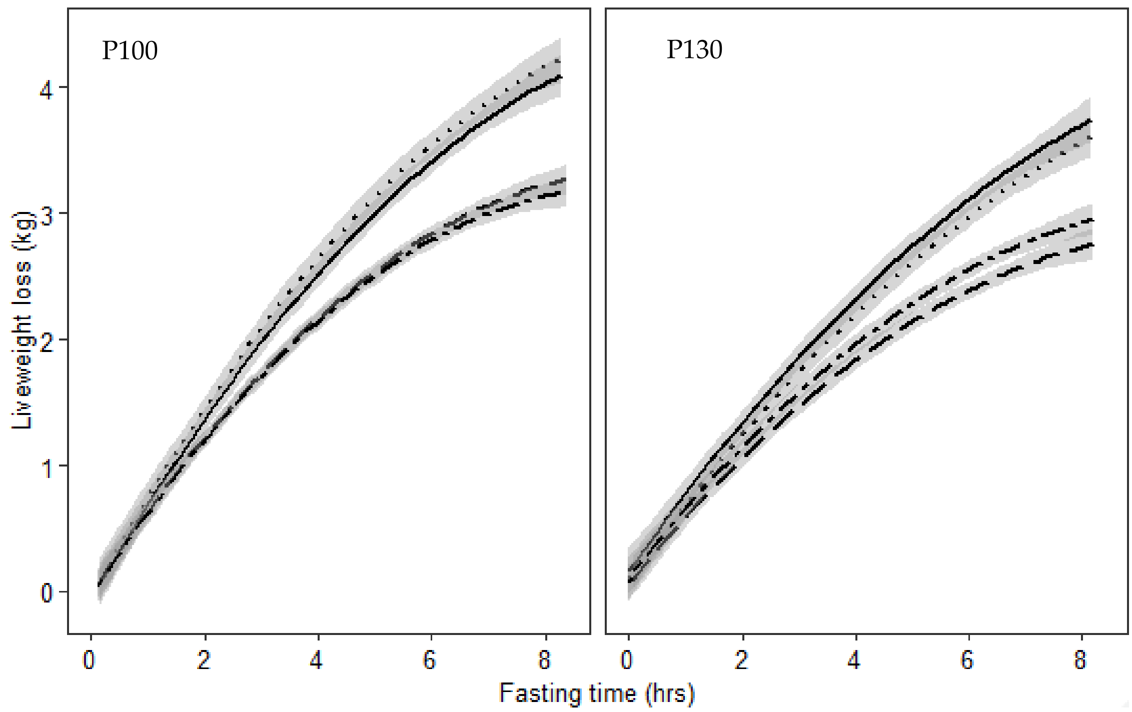


Figure 5.2 Change in live weight (with 95% CI, grey shade) of ewes (single: long dash line and twin: two dashed line) offered the Low_2 target range herbage availability and those (single: solid line and twin dotted line) offered the High herbage level at 100 days (P100) of pregnancy from the midpoint of a 17-day breeding period and 130 (P130) days. Herbage availability (Low_2 target range herbage availability: 900–1100 kg DM/ha; High ≥1400 kg DM/ha).

5.3.2 Validation stage

5.3.2.1 Herbage availability and proportion of live/green matter

The recorded average herbage masses were within the target ranges for all herbage availability levels and across studies except for the Low_1 herbage level (above) offered to non-pregnant ewes at Keeble farm (Appendix VI 2a). Overall, average herbage mass differed by availability level for all ewe studies ($p < 0.01$). In pregnant ewe studies, average herbage mass differed by farm and availability level ($p < 0.01$). Herbage mass was greater and more varied in the High than the Low_2 herbage level regardless of stage of pregnancy and farm. There was a significant herbage availability x farm interaction ($F_{1,16} = 10.12$, $p = 0.006$) in the P100 but not ($F_{2,18} = 3.23$, $p = 0.53$) in the P130 study. Average herbage mass offered to ewes also did not differ by stage of pregnancy ($p > 0.05$). In the pregnant ewe studies, average herbage masses offered in the High herbage level irrespective of stage of pregnancy varied greatly (1716 to 2226 kg DM/ha) for Keeble farm, Tuapaka farm (1712 − 2170) than Riverside farm (1442 to 1631 kg DM/ha) ($p < 0.05$). However, the masses offered under the Low_2 herbage level were comparable for all three farms ($p > 0.05$). Within farm, most herbage variability was observed in the High herbage level.

Overall, the proportion of herbage that was considered live (green) increased with herbage availability ($p < 0.05$). Further, the herbage levels offered to pregnant ewes had greater proportions ($p > 0.01$) of live or green herbage than that offered to non-pregnant ewes. Farm and/or farm x herbage availability did not ($p > 0.05$) affect the proportion of live matter.

5.3.2.2 Herbage chemical composition

The chemical composition of the herbage offered to non-pregnant ewes did not vary ($p > 0.05$) with herbage availability except for NDF ($p < 0.05$) (Appendix VI 2b). The herbage levels offered to non-pregnant ewes were also greater ($p < 0.05$) in DM, NDF and ADF, but were correspondingly lower ($p > 0.05$) in CP and ME compared with those offered to pregnant ewes. In the pregnant ewe study, all herbage chemical composition parameters varied with herbage availability and farm except for CP and ME. There was also a significant herbage availability x farm interaction ($p < 0.05$) for all herbage chemical composition parameters except for CP and ME. Stage of pregnancy and thus, time of year did not affect the chemical composition of herbage except for DM. The DM% of herbage increased with decreasing herbage level across ewe physiological status and farm ($p < 0.05$). There was variability in the rest of the herbage quality parameters with no predictable pattern across herbage availability and farm.

5.3.2.3 *Liveweight loss trends during calibration and validation stage*

Six-hour ewe liveweight change trends were compared for both calibration and validation stages. The regression equations derived after six hours of fasting during the validation stage are presented in Tables (5.3, 5.4 and 5.5). Overall liveweight loss varied among farms ($p < 0.001$) across ewe physiological state but not ($p > 0.05$) between herbage availability levels. The rate of liveweight loss differed ($p < 0.05$) between herbage levels and farm regardless of physiological state.

5.3.2.3.1 Non-pregnant ewe study

In the non-pregnant ewe study, the rate of liveweight loss was higher ($p < 0.05$) in both Medium and High herbage levels than the Low group (Table 5.3). Though not compared across farm, the rate of liveweight loss among the non-pregnant ewes offered the Medium herbage level was lowest at Riverside farm.

5.3.2.3.2 Pregnant ewe study

In the pregnant ewe study, all main effects were not significant ($p > 0.05$) except for time effects (first and second order fasting time polynomial) were significant ($p < 0.001$) (Tables 5.4 and 5.5). Further, among interactions, only, herbage availability x fasting time, stage of pregnancy x (first and second order fasting time polynomial), Study stage x fasting time and farm x fasting time were significant ($p < 0.05$).

5.3.2.3.2.1 P100 study

At P100, on both Keeble farm and Tuapaka farm, the rate of liveweight loss was greater for the High herbage level than the Low group regardless of pregnancy-rank (Table 5.4). Within pregnancy-rank, the rate of liveweight loss was comparable ($p > 0.05$) except for the Low herbage level on Tuapaka farm where the rate of liveweight loss was greater for twin than single-bearing ewes.

5.3.2.3.2.2 P130 study

At P130, on the Keeble farm, the rate of liveweight loss was comparable ($p > 0.05$) for all herbage levels across pregnancy-ranks (Table 5.5). On both Keeble farm and Riverside farm, ewes offered the Low_2 herbage level had comparable liveweight loss rate trends ($p > 0.05$). However, in ewes offered the High herbage level, the rate of liveweight loss was greater ($p < 0.05$) at Riverside farm than Keeble farm. The magnitude and rate of liveweight loss was comparable ($p > 0.05$) for Low_2 and High herbage levels at both Keeble farm and Tuapaka farm except ($p > 0.05$) for the twin-bearing ewes offered the High herbage level.

5.3.2.4 *Variability in liveweight loss at calibration and validation*

Results showed high variability in liveweight loss data (Tables 5.3, 5.4, 5.5). In non-pregnant ewes the greatest variability was observed (CV = 0.43−0.83) during the validation and least (CV = 0.39−78) observed during the calibration stage (Table 5.3). The greatest and most variable portion of variance explained by each model was observed (R^2 = 0.36−0.70) during the validation stage, and the least (R^2 = 0.33−0.65) for the calibration stage. Further, the highest CV% was consistently observed in the Low herbage availability level (CV = 51−78%) while the lowest CV was in the High group (CV = 39−43%) for both calibration and validation stages. The calibration stage of the non-pregnant ewe study conducted at Riverside farm using the Medium herbage level had the least liveweight loss rate and the most variable live weight data (CV% = 87%). Within the pregnant ewes, the liveweight loss rate within six hours of fasting was comparable ($p > 0.05$) for all herbage availability levels across stage of calibration except for the High.

In pregnant ewes, the greatest and most variable variability was observed (CV = 0.33−0.59) during the validation and least (CV = 0.33−37) observed during the calibration stage. However, the greatest and most variable portion of variance explained by each model was observed (R^2 = 0.47−0.78) during the validation stage, and the least (R^2 = 0.67−0.77) for the calibration stage (Tables: 5.4 and 5.5). The liveweight loss rate within six hours of fasting was comparable ($p > 0.05$) for all herbage availability levels across stage of calibration except for the High herbage level.

Table 5.3 Mean initial ("without delay") and final weight and prediction parameters with standard errors in parentheses, for non-pregnant ewe liveweight loss (kg) during a six-hour fasting period, by herbage availability level (Low, Medium, High) and farm (Keeble, Riverside). CV is the coefficient of variation and adjusted R^2 is a measure of goodness of fitness of the model. All models were significant at $p < 0.01$.

		Live weight (kg)		Predictor				
Farm	Availability level	Initial	Final	Intercept	Time	$Time^2$	Coefficient of variation (CV)	Adjusted R^2
*Keeble	Low	67.9(0.92)	65.3(0.92)	0.08(0.142)^{b}	0.25(0.045)^{a}	0.00(0.007)^{c}	0.78	0.33
	Medium	64.0(0.82)	62.6(0.82)	−0.11(0.142)^{b}	0.44(0.045)^{b}	−0.02(0.007)^{bc}	0.63	0.47
	High	65.3(0.72)	63.5(0.72)	−0.47(0.143)^{b}	0.68(0.046)^{c}	−0.04(0.007)^{b}	0.39	0.65
†Keeble	Low	65.9(1.07)	62.0(1.07)	−0.66(0.34)^{b}	0.53(0.111)^{bc}	−0.02(0.021)^{bc}	0.51	0.69
	Medium	66.2(1.31)	62.8(1.31)	−2.19(0.362)^{a}	0.80(0.121)^{d}	−0.15(0.022)^{a}	0.44	0.67
	High	67.0(1.23)	63.1(1.23)	−2.06(0.329)^{a}	0.85(0.107)^{d}	−0.13(0.02)^{a}	0.43	0.70
†Riverside	Medium	63.7(1.18)	62.8(1.18)	−0.07(0.070)	0.12(0.052)	0.10(0.008)	0.83	0.36

Live weight (Initial: live weight "without delay", Final: Live weight after eight hours of fasting). Asterisks *,† attached to farm name indicate the study stage dataset used for the analysis (*: calibration dataset, †: validation dataset). ^{abc}: different superscripts denote significant difference at $p < 0.05$ in within each column of predictors and farm. Availability level (Low herbage target range: 700–900 kg DM/ha, Medium: 1100 −1300, High: ≥ 1400). Model goodness of fit: the higher R^2 the better. All contrasts based on Sidak's multiple-comparisons tests.

Table 5.4 Mean initial ("without delay") and final weight and prediction parameters with standard errors in parentheses, of P100 ewe liveweight loss (kg) during a six-hour fasting periods, by herbage availability level (Low, High), farm (Keeble, Riverside) and pregnancy-rank (single, twin-bearing). CV is the coefficient of variation and adjusted R^2 is a measure of goodness of fitness of the model. All models were significant at $p < 0.01$.

			Live weight (kg)		Predictor				
Farm	TRT	PD	Initial	Final	Intercept	Time	Time2	Coefficient of variation (CV)	Adjusted R^2
*Riverside	Low	Single	64.8(0.65)	62.1(0.61)	−1.1(0.2)^{ab}	0.86(0.063)^{abc}	−0.066(0.009)^{abc}	0.33	0.73
	Low	Twin	70.0(0.76)	67.3(0.75)	−1.1(0.22)^{ab}	0.84(0.068)^{abc}	−0.064(0.01)^{abc}	0.29	0.77
	High	Single	69.7(0.69)	66.3(0.69)	−1.1(0.21)^{ab}	0.98(0.063)^{bc}	−0.067(0.01)^{abc}	0.34	0.72
	High	Twin	73.9(0.88)	70.6(0.87)	−0.9(0.21)^{ab}	0.90(0.063)^{abc}	−0.055(0.01)^{abc}	0.37	0.74
†Keeble	Low	Single	65.4(0.92)	63.6(0.88)	−1.0(0.27)^{ab}	0.67(0.083)^{ab}	−0.06(0.013)^{abc}	0.59	0.46
	Low	Twin	68.6(1.02)	66.8(1.00)	−1.1(0.27)^{ab}	0.69(0.082)^{ab}	−0.06(0.013)^{abc}	0.57	0.51
	High	Single	70.9(1.08)	67.3(1.06)	−1.6(0.27)^{a}	1.18(0.084)^{c}	−0.093(0.013)^{a}	0.36	0.72
	High	Twin	72.4(0.87)	69.1(0.86)	−1.5(0.27)^{a}	1.12(0.084)^{c}	−0.09(0.013)^{ab}	0.35	0.70
†Tuapaka	Low	Single	62.1(0.72)	59.1(0.67)	−0.3(0.26)^{b}	0.58(0.08)^{a}	−0.022(0.012)^{c}	0.48	0.58
	Low	Twin	67.6(0.63)	64.6(0.64)	−0.5(0.26)^{ab}	0.66(0.08)^{ab}	−0.034(0.012)^{bc}	0.43	0.61
	High	Single	67.3(1.09)	63.2(1.06)	−1.0(0.27)^{ab}	1.01(0.081)^{bc}	−0.068(0.012)^{abc}	0.33	0.72
	High	Twin	69.1(0.83)	65.2(0.78)	−1.0(0.26)^{ab}	0.99(0.08)^{bc}	−0.066(0.012)^{abc}	0.33	0.72

Live weight (Initial: live weight "without delay", Final: Live weight after eight hours of fasting). Asterisks *,† attached to farm name indicate the study stage dataset used for the analysis (*: calibration dataset, †: validation dataset). abc: different superscripts denote significant difference at $p < 0.05$ within each column of predictors. Availability level (Low herbage target range: 700–900 kg DM/ha, Medium: 1100−1300, High: ≥ 1400). Model goodness of fit: the higher R^2 the better. All contrasts based on Sidak's multiple-comparisons tests.

Table 5.5 Mean initial ("without delay") and final weight and prediction parameters with standard errors in parentheses, of P130 ewe liveweight loss (kg) during a six-hour fasting periods, by herbage availability (Low, High), farm (Keeble, Riverside) and pregnancy-rank (single, twin-bearing). CV is the coefficient of variation and adjusted R^2 is a measure of goodness of fitness of the model. All models were significant at $p < 0.01$.

			Live weight (kg)		Predictor			Coefficient of	Adjusted
Farm	TRT	PD	Initial	Final	Intercept	Time	$Time^2$	variation (CV)	R^2
*Keeble	Low	Single	69.9(0.88)	67.6(0.86)	−0.2(0.21)^{a}	0.50(0.068)^{a}	0.026(0.011)^{b}	0.38	0.67
	Low	Twin	72.9(0.69)	70.5(0.66)	−0.3(0.21)^{a}	0.53(0.068)^{ab}	0.032(0.011)^{ab}	0.37	0.67
	High	Single	72.5(0.91)	69.6(0.88)	0.1(0.21)^{a}	0.66(0.067)^{abc}	0.014(0.01)^{b}	0.37	0.70
	High	Twin	76.4(0.87)	73.4(0.83)	−0.6(0.2)^{a}	0.82(0.066)^{abc}	0.052(0.01)^{ab}	0.35	0.70
†Keeble	Low	Single	69.8(0.93)	68.0(0.91)	−0.6(0.19)^{a}	0.66(0.061)^{abc}	0.047(0.009)^{ab}	0.54	0.60
	Low	Twin	75.8(1.11)	74.1(1.10)	−0.7(0.19)^{a}	0.70(0.061)^{abc}	0.052(0.009)^{ab}	0.57	0.47
	High	Single	72.2(0.96)	68.8(0.88)	−0.8(0.2)^{a}	0.78(0.063)^{abc}	0.06(0.01)^{ab}	0.41	0.65
	High	Twin	77.7(1.01)	75.4(1.00)	−0.8(0.2)^{a}	0.81(0.064)^{abc}	0.065(0.01)^{ab}	0.39	0.64
†Riverside	Low	Single	69.1(0.89)	66.7(0.84)	−0.4(0.22)^{a}	0.63(0.07)^{abc}	0.039(0.011)^{ab}	0.45	0.65
	Low	Twin	74.7(0.87)	72.4(0.85)	−0.7(0.22)^{a}	0.74(0.071)^{abc}	0.056(0.011)^{ab}	0.41	0.68
	High	Single	73.2(0.88)	70.3(0.87)	−0.8(0.23)^{a}	0.97(0.075)^{c}	0.084(0.013)^{a}	0.32	0.78
	High	Twin	78.1(1.15)	75.4(1.13)	−0.8(0.23)^{a}	0.86(0.074)^{bc}	0.069(0.012)^{ab}	0.32	0.77
†Tuapaka	Low	Single	70.0(1.03)	67.2(1.00)	−0.5(0.21)^{a}	0.64(0.068)^{abc}	0.043(0.011)^{ab}	0.35	0.71
	Low	Twin	75.5(0.83)	72.7(0.81)	−0.5(0.22)^{a}	0.61(0.072)^{abc}	0.04(0.012)^{ab}	0.35	0.70
	High	Single	73.2(0.97)	69.5(0.94)	−0.5(0.22)^{a}	0.67(0.07)^{abc}	0.04(0.011)^{ab}	0.33	0.78
	High	Twin	76.7(0.87)	73.3(0.83)	−0.3(0.21)^{a}	0.58(0.069)^{ab}	0.027(0.011)^{ab}	0.36	0.74

Live weight (Initial: live weight "without delay", Final: Live weight after eight hours of fasting). Asterisks *,† attached to farm name indicate the study stage dataset used for the analysis (*: calibration dataset, †: validation dataset). abc: different superscripts denote significant difference at $p < 0.05$ within each column of predictors. Availability level (Low herbage target range: 700–900 kg DM/ha, Medium: 1100–1300, High: ≥1400). Model goodness of fit: the higher R^2 the better. All contrasts based on Sidak's multiple-comparisons tests.

5.3.2.5 Using correction equations can improve "without delay" ewe live weight estimation

5.3.2.5.1 Non-pregnant ewe study

The validated results showed that the ewe live weight correction equations for all feeding levels by feeding level and model developed in stage one, of the present study, predicted live weight with accuracy (Table 5.6, Figure 5.3) as shown by their low RPE and high r^2 and RPIQ values as compared with not using any correction method. The data presented in the Figure 5.3 is from Keeble farm. Riverside farm data is not presented on account of it being incomplete (only one availability level was evaluated).

At Keeble farm, compared with using the delayed live weights in ewes offered the Low herbage level, the specific equations to predict "without delay" live weight reduced error by 37.6% (0.91 kg) while using a mistaken equation (not meant for that herbage level) reduced error by 58.2% (1.38 kg). Within the ewes offered the Medium herbage level, using the specific equations to predict "without delay" live weight reduced error by 42.3% (1.28 kg) while using the mistaken equation reduced error by 40.4% (1.24 kg). Within the ewes offered the High herbage level, using the herbage-specific equations to predict "without delay" live weight reduced error by 57.6% (1.62 kg) while using the mistaken equation reduced error by 34% (0.95 kg) at Keeble farm. At Riverside farm, for the Medium herbage level, using the specific equations to predict "without delay" live weight increased error (introduced more error) by 28.1% (0.23 kg), and by 44.3% (0.36 kg) using the mistaken equation. The greatest accuracy was observed when estimating "without delay" live weight using the mistaken equations for the Low herbage level, both the mistaken and herbage level specific (separate) equation for the Medium herbage level and the specific equation for the High herbage level (Figure 5.3).

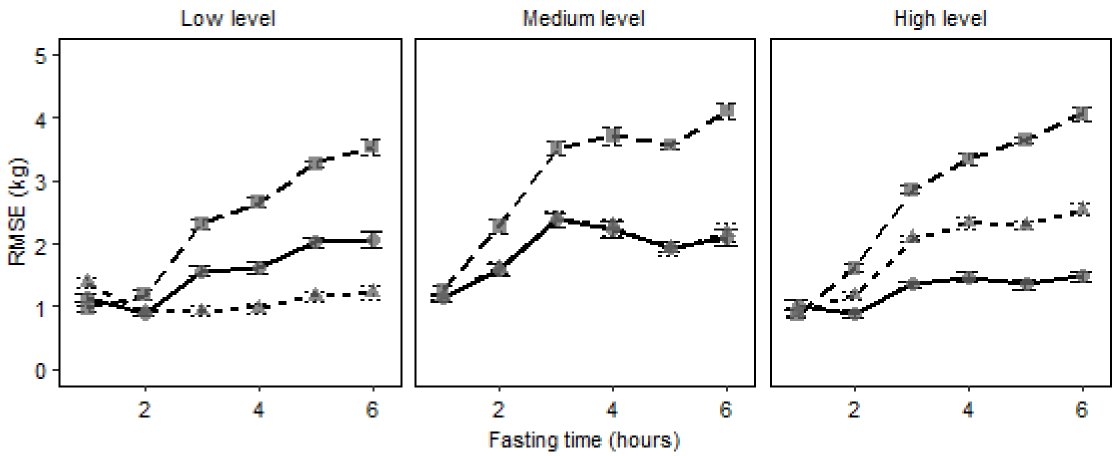


Figure 5.3 Change in root mean square error (RMSE) with associated standard deviation for the prediction of “without delay” live weight of non-pregnant ewes over fasting time when correction equations (dashed line with square points: no correction equation applied, dotted line with cross points: mistaken equation applied, solid line with circular points: availability level combination specific/separate equations applied) for each target herbage (Low, Medium and High) generated in stage one were applied on data collected in the summer season of 2020 at Keeble farm. Availability level (Low herbage target range: 700–900 kg DM/ha, Medium: 1100–1300 High: ≥1400). Correction equations: Herbage availability combination (separate: herbage specific equation correctly applied, mistaken any of the availability level specific equations wrongly applied to a different treatment combination).

Table 5.6 Measures of goodness of fit and accuracy (Bias, RMSE, RPE, RPD, RPIQ, r^2, CCC) of live weight ("without delay") prediction models (None: no model applied, separate: a separate/specific model applied, combined: pooled model where results were not significantly different and mistaken: model not developed for that availability level was applied) of non-pregnant ewes offered the Low_1, Medium or High herbage availability level during six hours of fasting tested on independent datasets (validation dataset) from Keeble farm and Riverside farm collected in 2020. Sample size (n) and weight ranges (kg).

			Live weight (kg)									
Farm	Herbage availability	Model	Actual Initial	Actual Final	Predicted Final	Bias	RMSE	RPE	RPD	RPIQ	r^2 %	CCC %
Keeble	Low_1	None				−1.85	2.39	3.67	3.00	4.47	98.7	99.3
		Separate	65.2(1.31)	61.8(1.34)	63.5(1.33)	−0.87	1.48	2.28	4.85	7.21	98.7	99.4
		Mistaken			64.7(1.34)	0.03	1.01	1.55	7.13	10.62	98.7	99.3
	Medium	None				−2.64	3.08	4.46	2.19	3.09	97.1	85.0
		Separate	66.3(1.23)	62.9(1.23)	64.4(1.22)	−1.23	1.80	2.61	3.74	5.27	97.1	99.2
		Mistaken			63.4(0.86)	−1.21	1.84	2.67	3.66	5.15	97.1	99.1
	High	None				−2.33	2.78	4.21	2.12	2.83	97.8	80.7
		Separate	65.9(1.07)	62.0(1.09)	64.9(1.09)	−0.51	1.16	1.76	5.08	6.77	97.8	99.0
		Mistaken			63.6(1.09)	−1.37	1.83	2.78	3.22	4.29	97.8	98.9
Riverside	Medium	None				−0.49	0.81	1.26	8.17	3.17	99.6	99.8
		Separate	63.7(1.18)	62.8(1.14)	65.0(1.13)	1.85	1.04	1.63	6.33	2.66	99.6	99.8
		Mistaken			65.1(0.80)	1.89	1.17	1.84	5.62	7.68	99.6	99.8

Herbage availability level (Low_1 target range herbage: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥ 1400 kg DM/ha). Interpretation of measures: The best model has the highest r^2, RPD, and RPIQ, and the lowest RMSE. Ranges for values: r^2 (0: indicates that the model explains none of the variability of the response data around its mean, 1.0 indicates that the model explains all the variability). RPD (< 1.4: weak, 1.4 < RPD < 2.0: reasonable, > 2.0: excellent). RPIQ (< 1.4: very poor, 1.4 < RPIQ < 1.7: fair, 1.7 < RPIQ < 2.0: good, 2.0 < RPIQ < 2.5: very good, > 2.5: excellent).

5.3.2.5.2 Pregnant ewe study

5.3.2.5.2.1 P100 study

The validated results showed that the ewe live weight correction equations for all feeding levels by feeding level, pregnancy-rank and model developed in stage one, predicted live weight with accuracy (Table 5.7, Figure 5.4) as shown by their low RPE and high r^2 and RPIQ values as compared with not using any correction method.

At Keeble farm, compared with using the delayed live weights in single-bearing ewes offered the Low herbage level, the herbage level specific equations to predict "without delay" live weight reduced error by 37% (0.64 kg), the combined equation by 32% (0.56 kg) while using mistaken equation (not meant for that herbage level) reduced error by 13 % (0.23 kg). Within the twin-bearing ewes, error was reduced by 39.1 % (0.61 kg), 41.0% (0.64 kg), 7.7% (0.12 kg) for the specific equation, combined and mistaken equations, respectively. The live weight estimation error was reduced more significantly among those ewes offered the High herbage level ($p < 0.01$). Single-bearing ewes offered the High herbage level, the specific equations to predict "without delay" live weight reduced error by 63% (1.87 kg), the combined equation by 63% (1.86 kg) while using mistaken equation reduced error by 55 % (1.63 kg). Within the twin-bearing, error was reduced by 61.0 % (1.68 kg), 63.0% (1.74 kg), 56% (1.54 kg) for the specific equation, combined and mistaken equations, respectively.

At Tuapaka farm, single-bearing ewes offered the Low herbage level, the specific equations to predict "without delay" live weight reduced error by 56% (1.25 kg), the combined equation by 56% (1.25 kg) while using mistaken equation reduced error by 47% (1.05 kg) compared with using the delayed live weights. Within the twin-bearing, error was reduced by 45 % (1.12 kg), 44.0% (1.11 kg), 40% (0.99 kg) for the specific, combined and mistaken equations, respectively. The estimation error reduction proportions were comparable for ewes offered both herbage levels ($p > 0.05$). Single-bearing ewes offered the High herbage level, the specific equations to predict "without delay" live weight reduced error by 53% (1.77 kg), the combined equation by 54% (1.78 kg) while using mistaken equation reduced error by 56 % (1.52 kg). Within the twin-bearing, error was reduced by 67 % (1.94 kg), 68% (1.97 kg), 58% (1.70 kg) for the specific, combined and mistaken equations, respectively.

Table 5.7 Measures of goodness of fit and accuracy (Bias, RMSE, RPE, RPD, RPIQ, r^2, CCC) of live weight ("without delay") prediction models (None: no model applied, separate: a separate/specific model applied, combined: pooled model where results were not significantly different and mistaken: model not developed for that availability level was applied) of ewes offered the Low_2, and High herbage levels by pregnancy-rank (PR) at 100 days of pregnancy (from the midpoint of a 17-day breeding period) and during six hours of fasting tested on independent datasets (validation dataset) from Keeble farm and Riverside farm collected in 2020. The range of values reflects the results of 30 random iterations of the models.

Farm	Herbage availability	Pregnancy-rank	Model	Live weight (kg) Actual Initial	Actual Final	Predicted Final	Bias	RMSE	RPE%	RPD	RPIQ	r^2%	CCC%
Keeble	Low_2	Single	None				−1.47	1.74	2.67	3.69	5.75	98.3	96.9
			Separate	65.4	63.6	65.2(1.53)	0.45	1.10	1.68	5.83	9.09	98.3	90.2
			Combined			65.0(1.53)	0.32	1.18	1.81	5.44	8.47	98.3	94.9
			Mistaken			66.1(1.53)	0.94	1.51	2.30	4.28	6.67	98.3	84.6
	Low_2	Twin	None				−1.28	1.56	2.27	4.55	7.50	98.9	99.1
			Separate	68.6	66.8	69.6(2.1)	0.56	0.95	1.38	7.56	12.32	98.9	89.9
			Combined			69.5(2.1)	0.51	0.92	1.34	7.81	12.72	98.9	94.8
			Mistaken			70.6(2.1)	1.13	1.44	2.10	4.99	8.13	98.9	86.7
	High	Single	None				−2.62	2.97	4.19	2.57	2.79	97.8	98.2
			Separate	70.9	67.3	71.0(1.06)	−0.19	1.10	1.55	6.94	7.55	97.8	76.7
			Combined			71.1(1.06)	−0.14	1.11	1.55	6.94	7.55	97.8	99.9
			Mistaken			70.0(1.06)	−0.78	1.34	1.89	5.70	6.19	97.8	96.8
	High	Twin	None				−2.67	2.76	3.81	2.19	3.53	97.6	95.5
			Separate	72.4	69.1	72.5(2.84)	−0.39	1.08	1.49	5.6	8.94	97.6	64.7
			Combined			72.9(2.84)	−0.19	1.02	1.41	5.93	9.46	97.6	92.6
			Mistaken			71.7(2.84)	−0.83	1.22	1.68	4.96	7.91	97.6	99.4
Tuapaka	Low_2	Single	None				−1.92	2.25	3.60	2.26	2.71	96.4	86.1
			Separate	62.1	59.1	62.0(1.67)	0.08	1.00	1.57	5.10	6.11	96.4	83.8
			Combined			61.8(1.67)	−0.05	1.00	1.60	5.11	6.12	96.4	97.8
			Mistaken			63.0(1.67)	0.60	1.20	1.90	4.24	5.08	96.4	84
	Low_2	Twin	None				−1.97	2.50	3.72	1.97	2.48	91.7	99.9
			Separate	67.6	64.6	67.7(2.7)	−0.05	1.38	2.01	3.56	4.52	91.7	58.8
			Combined			67.6(2.7)	−0.1	1.39	2.03	3.55	4.52	91.7	97.8
			Mistaken			68.8(2.7)	0.55	1.51	2.21	3.26	4.13	91.7	98.8
	High	Single	None				−2.83	3.33	4.94	2.28	2.9	96.1	99.9
			Separate	67.3	63.2	66.8(1.04)	−0.37	1.56	2.26	4.86	6.18	96.1	77.9
			Combined			67.0(1.04)	−0.32	1.55	2.31	4.89	6.22	96.1	99.8

		Mistaken			65.8(1.04)	−0.96	1.81	2.70	4.19	5.33	96.1	97.8
High	Twin	None	69.1	65.2		−2.64	2.91	4.18	2.02	2.32	97.8	99.7
		Separate			68.6(2.78)	−0.33	0.97	1.36	6.03	6.93	97.8	63.4
		Combined			69.0(2.78)	−0.13	0.94	1.38	6.28	7.22	97.8	85.7
		Mistaken			67.9(2.78)	−0.77	1.21	1.69	4.85	5.57	97.8	97.6

Herbage availability (Low_2 target range herbage: 900–1100 kg DM/ha, High: ≥ 1400 kg DM/ha). Pregnancy-rank (PR: 1=single-bearing, 2=twin-bearing). Interpretation of measures: The best model has the highest r^2, RPD, and RPIQ, and the lowest RMSE. Ranges for values: r^2 (0: indicates that the model explains none of the variability of the response data around its mean, 1.0 indicates that the model explains all the variability). RPD (< 1.4: weak, 1.4 < RPD < 2.0: reasonable, > 2.0: excellent). RPIQ (< 1.4: very poor, 1.4 < RPIQ < 1.7: fair, 1.7 < RPIQ < 2.0: good, 2.0 < RPIQ < 2.5: very good, > 2.5: excellent).

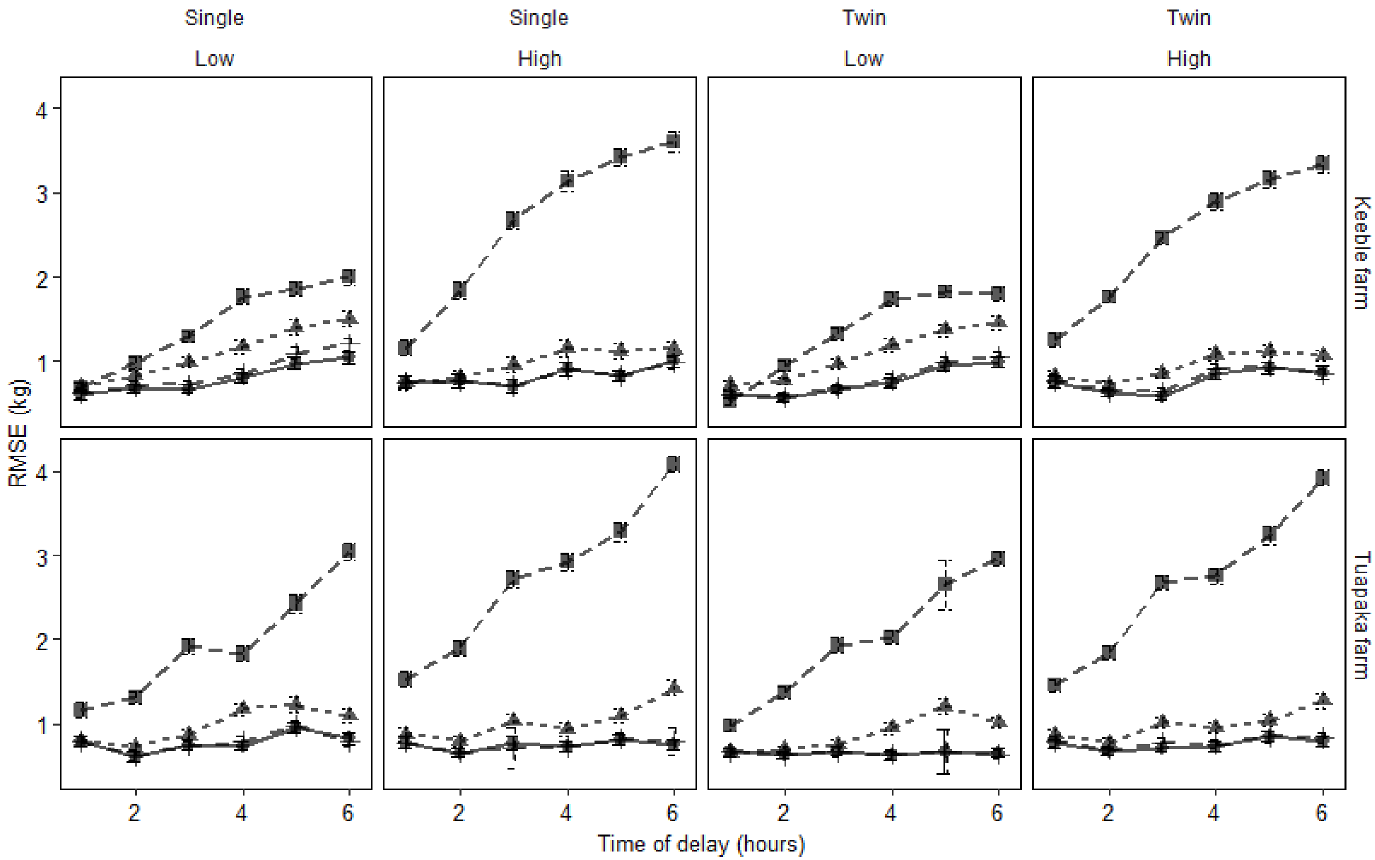


Figure 5.4 Change in root mean square error (RMSE) with associated standard deviation for the prediction of "without delay" live weight of P100 ewes over fasting time when specific correction equations (dashed line with square points: no correction equation applied, dash dotted line with cross points: availability level combination specific equations, solid line with circular points: combined equations and dotted line with triangular points: mistaken equations) for each target herbage level (Low_2 and High) and pregnancy-rank (single and Twin) generated in stage one were applied on data collected in the winter season of 2020 by Farm (Tuapaka, Riverside). Availability level (Low_2 herbage target range: 900–1100 kg DM/ha, High: ≥1400). Herbage availability combination correction equation (Separate: herbage, stage of pregnancy and pregnancy-rank specific, combined: consolidated equations with similar effect, mistaken any of the availability level specific equations wrongly applied to a different treatment combination).

5.3.2.5.2.2 P130 study

The validated results showed that the ewe live weight correction equations for all feeding levels by feeding level, pregnancy-rank and model developed, predicted live weight with substantial accuracy (Table 5.8, Figure 5.5) as shown by their low RPE and high r^2 and RPIQ values as compared with not using any correction method. The reduction in the proportion of the live weight prediction error for ewes offered the Low or High herbage level was greatest at Riverside farm regardless of pregnancy-rank ($p < 0.05$). Although the proportion of the live weight prediction was comparable ($p > 0.05$) among ewes offered the Low herbage level, Tuapaka had greater error for ewes offered the High herbage level than Keeble farm ($p < 0.05$).

At Keeble farm, compared with using the delayed live weights in single-bearing ewes offered the Low herbage level, the herbage-specific equations to predict "without delay" live weight reduced error by 43.3% (0.68 kg), the combined equation by 42.0% (0.66 kg) while using mistaken equation reduced error by 27.4% (0.43 kg). Within the twin-bearing, error was reduced by 36.7% (0.61 kg), 38.6% (0.64 kg), 26.5% (0.44 kg) for the specific equation, combined and mistaken equations, respectively. The estimation error was reduced more significantly among those ewes offered the High herbage level ($p < 0.01$). Among single-bearing ewes offered the High herbage level, the specific equations to predict "without delay" live weight reduced error by 60.3% (1.26 kg), the combined equation by 59.8% (1.25 kg) while using mistaken equation reduced error by 62.7 % (1.31 kg). Within the twin-bearing, error was reduced by 50.2% (1.14 kg), 50.7% (1.15 kg), 52.4% (1.19 kg) for the specific equation, combined and mistaken equations, respectively.

At Tuapaka farm, single-bearing ewes offered the Low herbage level, the specific equations to predict "without delay" live weight reduced error by 56.0% (1.17 kg), the combined equation by 56.0% (1.17 kg) while using mistaken equation reduced error by 53.6% (1.12 kg) compared with using the delayed live weights. Within the twin-bearing, error was reduced by 58.5% (1.24 kg), 58.5% (1.24 kg), 57.5% (1.22 kg) for the specific, combined and mistaken equations, respectively. The magnitude of the decrease in the live weight prediction error was comparable for ewes offered both the Low_2 and High herbage levels ($p > 0.05$). In single-bearing ewes offered the High herbage level, the specific equations to predict "without delay" live weight reduced error by 61.6% (1.72 kg), the combined equation by 61.6% (1.72 kg) while using mistaken equation reduced error by 54.8 % (1.53 kg). Within the twin-bearing, error was reduced by 60.9 % (1.54 kg), 60.5% (1.53 kg), 58.9% (1.49 kg) for the specific, combined and mistaken equations, respectively.

Among single-bearing ewes offered the Low herbage level at Riverside farm, the specific equations to predict “without delay” live weight reduced error by 65.8% (1.25 kg), the combined equation by 65.8% (1.25 kg) while using mistaken equation reduced error by 61.1% (1.16 kg) compared with using the delayed live weights. Within the twin-bearing, error was reduced by 68.2% (1.22 kg), 68.7% (1.23 kg), 60.3% (1.08 kg) for the specific, combined and mistaken equations, respectively. The accuracy in predicting the “without delay” live weights were comparable for ewes offered both herbage levels ($p < 0.01$). In single-bearing ewes offered the High herbage level, the specific equations to predict “without delay” live weight reduced error by 69.0% (1.45 kg), the combined equation by 68.6% (1.44 kg) while using mistaken equation reduced error by 69.0 % (1.33 kg). Within the twin-bearing, error was reduced by 62.7% (1.21 kg), 64.2% (1.24 kg), 68.9% (1.33 kg) for the specific, combined and mistaken equations, respectively.

Table 5.8 Measures of goodness of fit and accuracy (Bias, RMSE, RPE, RPD, RPIQ, r^2, CCC) of live weight ("without delay") prediction models (None: no model applied, separate: a separate/specific model applied, combined: pooled model where results were not significantly different and mistaken: model not developed for that availability level was applied) of ewes offered the Low_2, and High herbage levels by pregnancy-rank (PR) at 130 days of pregnancy and during six hours of fasting tested on independent datasets (validation dataset) from Keeble farm, Tuapaka farm and Riverside farm collected in 2020. Sample size (n) and weight ranges (kg).

Farm	Herbage availability	Pregnancy-rank	Model	Live weight (kg) Actual Initial	Actual Final	Predicted Final	Bias	RMSE	RPE %	RPD	RPIQ	r^2 %	CCC %
Keeble	Low_2	Single	None				−1.29	1.57	2.24	4.18	6.07	98.6	95.8
			Separate	69.8(0.93)	68(0.91)	70.5(0.91)	0.36	0.89	1.27	7.37	10.69	95	96.9
			Combined			70.5(0.91)	0.41	0.91	1.31	7.17	10.4	98.6	98.6
			Mistaken			71.1(0.91)	0.74	1.14	1.64	5.74	8.33	98.6	97.3
		Twin	None				−1.34	1.66	2.19	4.74	8.06	98.6	97
			Separate	75.8(1.11)	74.1(1.1)	76.7(1.1)	0.42	1.05	1.38	7.51	12.76	82.6	90.3
			Combined			76.6(1.1)	0.36	1.02	1.35	7.7	13.08	98.6	98.8
			Mistaken			77.2(1.1)	0.69	1.22	1.61	6.45	10.97	98.6	97.9
Keeble	High	Single	None				−1.88	2.09	2.9	3.24	4.65	98.8	87.8
			Separate	72.2(1.36)	68.8(1.25)	72(1.25)	0.25	0.83	1.15	8.14	11.73	98.8	99
			Combined			72(1.25)	0.26	0.84	1.16	8.11	11.66	98.8	99
			Mistaken			71.4(1.25)	−0.09	0.78	1.08	8.72	12.55	98.8	98.4
		Twin	None				−1.96	2.27	2.92	3.15	4.82	97.7	93.8
			Separate	77.7(1.02)	75.4(1.02)	77(1.86)	0.26	1.13	1.46	6.3	9.66	83.9	76.7
			Combined			78.5(1.01)	0.21	1.12	1.44	6.38	9.79	97.8	98.4
			Mistaken			77.9(1.01)	−0.15	1.08	1.39	6.6	10.13	97.8	98.8
Tuapaka	Low_2	Single	None				−1.77	2.09	2.98	3.48	2.85	98.4	92.3
			Separate	70.0(1.02)	67.2(1.00)	69.7(1.00)	−0.05	0.92	1.31	7.89	6.47	98.4	99.1
			Combined			69.8(1.00)	0.03	0.92	1	1.31	7.92	98.4	98.4
			Mistaken			70.3(1.03)	0.35	0.97	1.13	1.39	7.44	98.4	98.4
		Twin	None				−2.14	2.12	2.81	2.77	3.23	97.8	88.4
			Separate	75.5(0.83)	72.7(0.81)	75.3(0.81)	−0.32	0.88	1.16	6.71	7.82	79.7	89.2
			Combined			75.2(0.82)	−0.38	0.88	0.9	1.17	6.65	79.7	79.7
			Mistaken			75.9(0.82)	−0.03	0.9	0.94	1.2	6.51	79.6	79.6
Tuapaka	High	Single	None				−2.53	2.79	3.82	2.39	3.13	97.9	85.6
			Separate	73.2(0.96)	69.5(0.94)	72.7(0.93)	−0.43	1.07	1.46	6.25	8.17	97.8	98.5
			Combined			72.7(0.94)	−0.42	1.07	1.06	1.46	6.27	97.8	97.8

			Mistaken			72.1(0.94)	−0.77	1.26	1.31	1.72	5.3	97.8	97.8
		Twin	None				−2.58	2.53	3.3	2.37	2.63	97.4	84.4
			Separate	76.7(0.87)	73.3(0.82)	76.5(0.83)	−0.44	0.99	1.29	6.05	6.71	78.7	88.6
			Combined			76.4(0.83)	−0.49	1	0.94	1.31	5.99	78.9	78.9
			Mistaken			75.9(0.84)	−0.83	1.04	1.09	1.36	5.76	78.7	78.7
Riverside	Low_2	Single	None				−1.69	1.9	2.47	2.75	3.37	99.1	99.1
			Separate	69.1(0.89)	66.7(0.84)	69.1(0.84)	−0.06	0.65	0.69	0.94	9.86	99.2	99.2
			Combined			69.2(0.84)	−0.01	0.65	0.68	0.94	9.86	99.2	99.2
			Mistaken			69.7(0.84)	0.32	0.74	0.8	1.07	8.66	99.2	99.2
		Twin	None				−1.61	1.79	2.16	2.4	3.31	99.2	99.2
			Separate	74.7(0.87)	72.4(0.85)	75.0(0.85)	0.13	0.57	0.6	0.76	10.4	99.2	99.2
			Combined			74.9(0.85)	0.08	0.56	0.59	0.75	10.59	99.2	99.2
			Mistaken			75.5(0.85)	0.41	0.71	0.76	0.95	8.35	99.2	99.2
	High	Single	None				−1.85	2.1	2.54	2.87	3.01	99	98.9.
			Separate	73.2(0.88)	70.3(0.87)	73.3(0.87)	0.21	0.65	0.7	0.89	9.72	99	98.9
			Combined			73.4(0.87)	0.22	0.66	0.71	0.9	9.58	99	98.9
			Mistaken			72.8(0.87)	−0.12	0.65	0.67	0.89	9.72	99	98.9
		Twin	None				−1.69	1.93	2.18	2.47	4.28	99.5	99.5
			Separate	78.1(1.14)	75.4(1.12)	78.5(1.13)	0.42	0.72	0.78	0.92	11.49	99.5	99.5
			Combined			78.4(1.13)	0.37	0.69	0.75	0.88	11.99	99.5	99.5
			Mistaken			77.9(1.13)	0.03	0.6	0.61	0.77	13.78	99.5	99.5

Availability level (Low_2 target range herbage availability: 900–1100 kg DM/ha, High: ≥1400 kg DM/ha). Pregnancy-rank (PR: 1=single-bearing, 2=twin-bearing). Interpretation of measures: The best model has the highest r^2, RPD, and RPIQ, and the lowest RMSE. Ranges for values: r^2 (0: indicates that the model explains none of the variability of the response data around its mean, 1.0 indicates that the model explains all the variability). RPD (< 1.4: weak, 1.4 < RPD < 2.0: reasonable, > 2.0: excellent). RPIQ (< 1.4: very poor, 1.4 < RPIQ < 1.7: fair, 1.7 < RPIQ < 2.0: good, 2.0 < RPIQ < 2.5: very good, > 2.5: excellent).

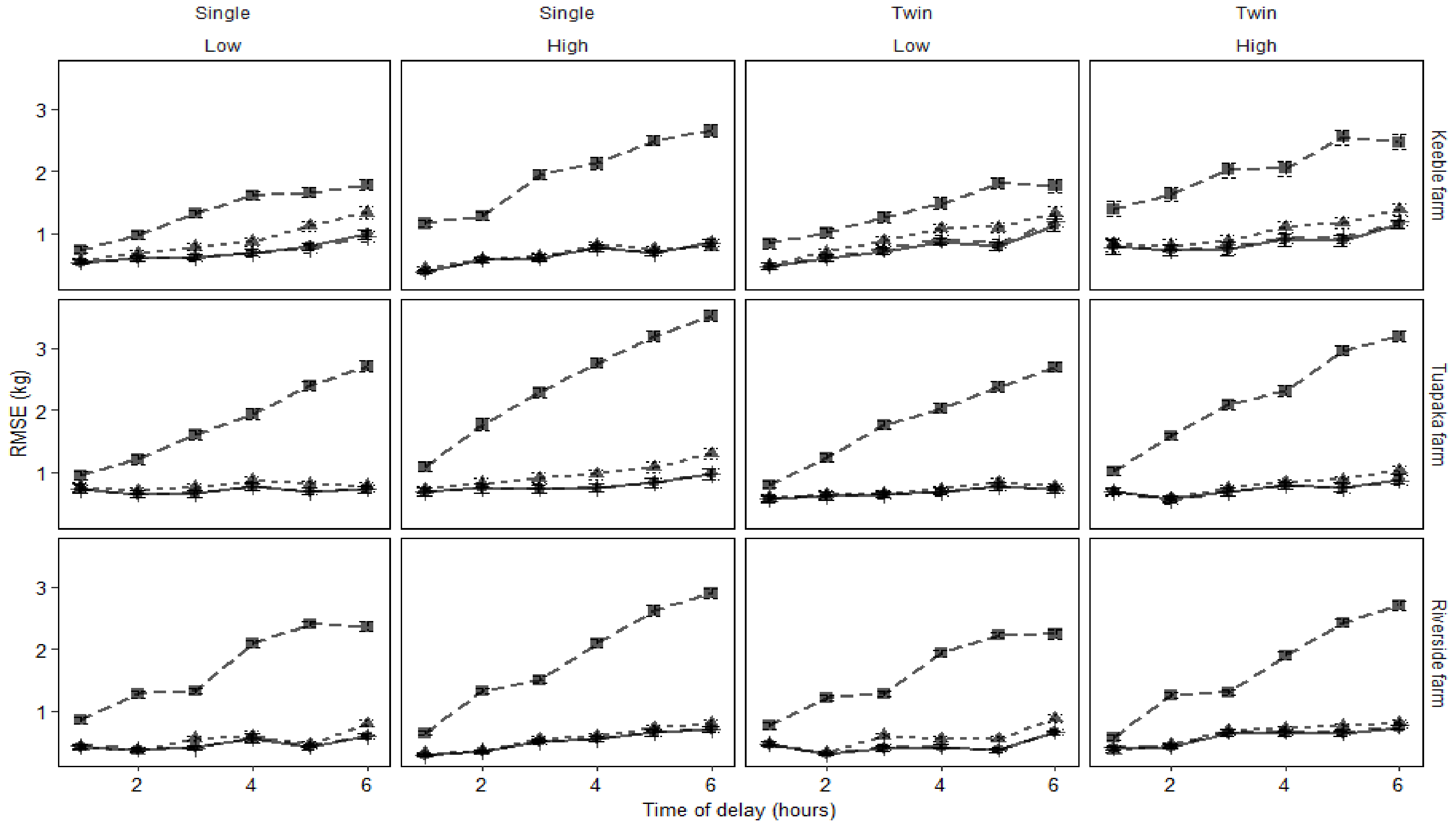


Figure 5.5 Change in root mean square error (RMSE) with associated standard deviation for the prediction of “without delay” live weight of P130 ewes over fasting time when specific correction equations (dashed line with square points: no correction equation applied, dash dotted line with cross points: treatment combination specific equations, solid line with circular points: combined equations and dotted line with triangular points: mistaken equations) for each target herbage (Low and High) and pregnancy-rank (single and Twin) generated in stage one were applied on data collected in the winter season of 2020 by farm (Keeble, Tuapaka, Riverside). Availability level (Low2 herbage target range: 900–1100 kg DM/ha, High: ≥1400). Correction equations: Treatment combination specific: herbage, stage of pregnancy and pregnancy-rank specific, combined: consolidated equations with similar effect, mistaken any of the availability level specific equations wrongly applied to a different treatment combination).

5.4 Discussion

The aim of the present study was to determine the effect of herbage availability, physiological status of a ewe on the rate of liveweight loss in ewes during fasting and determine if "without delay" live weight of ewes can be predicted with substantial accuracy (reduced error of prediction) from delayed weight.

5.4.1 Calibration stage

The findings indicated that overall, the ewes lost live weight (non-pregnant: 2.4−3.6 kg (3.8−5.3% of initial ewe live weight); P100: 3.1.0−5.0.0 (4.9−6.9% of initial ewe live weight); P130: 2.8−3.5 kg (3.8−4.7% of initial ewe live weight) between each weighing throughout the fasting period. The magnitude of this weight change is likely to influence the reliability of live weight measures which may have implications for management decisions on-farm and for research unless it can be corrected for. The current study indicates that the rate of liveweight loss was affected by both feeding herbage availability and physiological state suggesting that different equations may be required to correct for liveweight loss across herbage availability levels and physiological state during a fasting event. Similarly, in ewe lambs, the rate of liveweight loss was found to be influenced by feeding level (Chapter 4).

The variation in ewe liveweight loss by herbage availability was likely due to differences in gut-fill volume resulting from differences in feed composition notably dry matter (DM%) content of the herbage. The amount of DM% contained in the herbage was consistently lowest (highest moisture content) in the High herbage level and highest in the Low herbage level. It appears that the ewes were consuming more water from the High herbage levels than the Low herbage. The effect of DM% was greater among the pregnant ewes (study conducted during winter) than non-pregnant ewe (summer). Season affects the chemical composition of herbage which explains the higher herbage dry matter DM% and thus, relatively lower liveweight loss rate in summer than winter. The current studies were carried out in different seasons of the year and thus, different ambient temperature. However, since different physiological states were studied in different seasons, it was hard to separate the confounding effect of both season and physiological state. Exposure to colder temperatures has been reported to increase the reticulo-rumen motility, the passage rate of gut particles and to reduce the gut-fill retention time (Kennedy, 1985; Bernabucci et al., 1999). The ambient temperature in summer when the non-pregnant ewe trial was conducted, was higher than that for winter when the pregnant ewe trial was conducted which might explain the higher weight loss in the later.

Though not directly comparable, the high dry matter % (47.4%) offered to non-pregnant ewes and the low dry matter (15.7–19.6%) at 100 days in pregnancy and (15.9–18.8%) at 130 days in pregnancy in both High and Low herbage levels, respectively, meant they were likely to have consumed differing amounts of water in their herbage intake. Ewes need to consume approximately 3.0% of their live weight (Lloyd et al., 1978; McDonald, 2002). This suggests that in the current study a 66.4 kg ewe offered the any of the herbage levels, would have consumed an average of approximately 1.99 (0.03 x 66.4) kg DM. If non-pregnant and offered Low1 herbage level (47.4% DM), it would have consumed 2.20 (1.99 x 1.10) litres of water. Similarly, the same non-pregnant ewe offered the High herbage level (34.4% DM) would have consumed 3.8 (1.99 x1.91) litres of water. At P100, offered the Low_2 herbage level (19.6% DM), the same ewe would have consumed 8.2 (1.99 x 4.1) litres of water. Similarly, offered the High herbage level (15.7% DM), the same ewe, would have consumed 10.7 (1.99 x 5.4) litres of water. This extra water in herbage would be excreted faster through urine than herbage via faeces. The findings, therefore, suggest that the greater the DM% content of herbage the lower the rate of ewe liveweight loss and vice versa. Further, the greater proportions of CP and ME, but with correspondingly lower fibre (DM, NDF and ADF) may have been responsible for the lower rate of ewe liveweight loss in summer compared with winter. Dry matter and fibre have been reported to increase with herbage density (Toupet et al., 2020). Higher levels of fibre increase water holding capacity of the gut and thus the rumen clearance. The proportion of structural carbohydrates responsible (fibre) increases and that of fermentable carbohydrates and pectins decreases in drier seasons while in wet seasons the reverse is true (Crampton and Jackson, 1944; Litherland et al., 2002; Warly et al., 2004; Särkijärvi et al., 2012; Mir and Ahmed, 2017; Ekanayake et al., 2019). Therefore, it is not surprising that herbage in the non-pregnant ewe study (summer) had greater DM% and fibre, and thus, the lower rate of ewe liveweight loss than pregnant ewe study (winter). Greater structural carbohydrate results in a decrease in the fractional rate of fluid passing through the rumen thereby increasing the water holding capacity of the gut.

The results indicated that the rate of liveweight loss was influenced by the physiological state (pregnancy status and stage of the pregnancy) but, not pregnancy-rank. The results support the liveweight losses (3.8 to 5.0% and 4.0 to 7.2%) reported by Hughes (1976) and Burnham et al. (2009) in two-tooth and mature pregnant ewes respectively. The results from our study further indicated that ewe liveweight loss at day 100 was greater that at day 130. This finding is corroborated by Burnham et al. (2009) who reported greater ewe liveweight loss at day 70 of pregnancy than at day 130 (9.8 vs. 7.5%). The greater liveweight loss at day 100 may be due to a relatively smaller uterus volume compared with day 130, resulting in less constraint on the rumen volume. Thus, the clearance of a larger rumen volume at day 100 having a greater effect on overall liveweight loss. A negative relationship has been reported between rumen volume and uterus volume in pregnant ewes between day 72 and day 144 of gestation (Forbes, 1969). The authors reported a rumen volume decrease of 3.6 litres and uterus volume increase of 2.8 litres at days 72 and 144 respectively. The results, therefore, suggest that the rate of liveweight loss appears to decrease with advancing pregnancy. The finding that pregnancy-rank did not affect the rate of ewe liveweight loss, contrasts Burnham et al. (2009) who reported a greater proportional liveweight loss in single than twin-bearing ewes at day 130 of pregnancy. This discrepancy warrants further investigations. It appears pregnancy stage is likely more relevant in ewe liveweight loss than having single or twin foetuses. This might be attributed to greater energy needs that come with changes in pregnancy stage (days in pregnancy) compared with number of foetuses carried especially in early stages of pregnancy. Further, the observed differences in liveweight loss due to stages of pregnancy were not unexpected. The gestation period of a sheep is 147 days. The last trimester of gestation is the period of rapid conceptus growth (which includes foetus(es), fluid and placenta). Therefore, rapid changes in total weights are observed especially when one considers in these breed types the conceptus mass at term can be 16 to 18 kg in total weight(Kenyon et al., 2007a). Equations such as those by Gomptez (Freer et al., 2007) show just how exponential the foetal weight gain in this period is. Feeding guidelines clearly state that this is the period of rapid increase in feed demand to meet the nutritional increases required. Thus it is important for farmers to be able to determine if feeding levels are meeting the expected feeding requirements, allowing total weight of the ewe to increase with expected gains of the conceptus mass (i.e. so that she does not have to draw on her own body reserves significantly to meet this increased demand).

The current study utilized mixed-aged ewes of 3 – 5 years. A ewe reaches maturity at 3 years after which age effect becomes minimal (Cake et al., 2006; Semakula et al., 2020a). Therefore, differences in age in the current study ewes were, not expected to affect the liveweight loss rate. There was in-flock and between-flock liveweight loss variation in the ewes used. The in-flock variations were comparable across farms. The in-flock variations were accounted for as random variability while the between-flock variations were captured under farm effect in the linear mixed effects model. Further, some individual animals remained highly mobile during weighing. TruTest weighing scales have an algorithm that can quickly stabilize weight measurements even in highly mobile animals. Therefore, it is unlikely that the scale accuracy was affected, hence impacting the findings. The findings indicated farms differences which could have affected the findings as stated in the discussion section.

5.4.2 Validation stage

The significant polynomial regression between liveweight loss and time off feed and the subsequent linear association between delayed and “Without delay live weight” supports the concept of the relationship between weight loss and “Without delay” live weight. This is premised on the hypothesis that the amount of weight lost per unit time varies depending on herbage availability. It was observed that the weight prediction equations became more curvilinear than linear when herbage DM% decreased or as herbage availability was increased. Similar observations were made in a study with ewe lambs (Chapter 4).

A comparison of liveweight loss trends using calibration and validation datasets demonstrated significant differences in overall liveweight loss between farms. The results also demonstrated significant liveweight loss rates between herbage levels and farms. Further, the results indicated high CV % associated with this liveweight loss, which was highest at Keeble farm and lowest at Riverside farm. These finding point to potential differences that may have existed between sites. Notably, the herbage target ranges varied in herbage levels and dry matter content which might explain the differential weight losses on different farms. Additionally, at both Keeble farms and Tuapaka farm, live weights were recorded manually by the operator whereas at Riverside farm, weights were automatically recorded. Comparison weighing was done using two 20 kg loads at the start of each weighing event and therefore, differences cannot be due to starting calibration error. However, an automated weighing system regularly readjusts the scale to zero, thereby reducing the error introduced due to shifts in the position of the crate, this does not occur in manual systems.

Preferably, weighing without any delay (immediately off pasture) should provide ewe live weight measurements with least error. However, if this is not achievable, the validating process has demonstrated that correction equations can be used to supply corrected live weights (cW_0) that are more accurate estimates of the “without delay” live weight(aW_0) than a delayed live weight (dW_t). This highlights a major step towards achieving improved (precise) live weight measurement in sheep production.

The accuracy of the correction equations was significantly impacted by herbage availability, physiological state of a ewe, stage of the pregnancy-rank, the period of delay in recording the weight and farm. This supports our previous findings in Chapter 4, in which we found significant effects of herbage availability and season on the rate of liveweight loss of ewe lambs. Further, the results are in partial agreement with Wishart et al. (2017) who reported a significant impact of grazing location on precision of mature ewe live weight correction equation but, not time of delay. As expected, the authors showed that the precision of the correction equations was affected by the factors associated with fluctuations in gut-fill (Coates and Penning, 2000b; Wishart et al., 2017).

The correction equations had comparable stability/robustness overtime when predicting “without delay” live weight from the delayed live weights. The accuracy of ewe live weight correcting equations was greater for High herbage level than Low herbage level. This contrast with our previous findings (Chapter 4) where we reported more equation stability when predicting “without delay” live weight in ewe lambs offered the Low diet than the Medium or High herbage level. The lower quantity (kg DM/ha) of grazing herbage for the Low herbage level could have restricted the gut-fill thereby eliciting a response to reduce ruminal emptying. In addition, ewes offered the High herbage level had access to more variable herbage ranges (1500−2100 kg DM/ha) than those offered the Low herbage level (700–900 kg DM/ha for non-pregnant and 1000−1200 kg DM/ha for pregnant ewes) which might explain their associated greater error rates. However, it has previously been reported that intakes do not increase above the herbage level of ≈ 1400 kg DM/ha (Morris and Kenyon, 2004) and thus, this potential explanation does not hold.

In the validation study, we switched the correction equations applying them to mismatching ewe live weights and/or applied them on consolidated datasets regardless of study farm. Ideally, the greatest accuracy of “without delay” live weight prediction would be expected when herbage specific/separate equations were applied to delayed data. It is not clear why in the non-pregnant ewe study the mistaken equations gave greater accuracy than using a herbage availability level specific equation in ewes offered the Low herbage level or comparable accuracies for the Medium herbage level. Further, the results suggest that applying an equation from a different herbage level, stage of pregnancy or pregnancy-rank to predict the “without delay” live weight from delayed live weight would be a better option than using the delayed weights themselves. Further, applying the correction equations on consolidated rather than farm-specific datasets yielded mixed results, with greater, comparable or lesser live weight accuracies. The validations were conducted using a range of herbage availability levels and live weights which should cover most situations in an extensive sheep rearing system grazing a ryegrass-based diet. The use of simple and multiple linear regression equations based on time stamps to predict liveweight loss and to predict “without delay” live weight in ewe lambs has been previously reported in Chapter 4. In Chapter 4, the “without delay” live weight was predicted based herbage availability, season, and time off herbage with and supplied data on the levels of accuracy the equations had compared with not using the equation. The current study supports the hypothesis in Chapter 4 and by Wishart et al. (2017) that quality and quantity of herbage, physiological state of a ewe as well as environmental factors such farm and grazing location (Hughes, 1976; Moyo and Nsahlai, 2018) impact liveweight variation. These factors may be interact causing the differences between ewes from different farming/grazing locations, physiological state and feeding levels.

The results of the present study demonstrated that it is possible to obtain substantially accurate estimates of “without delay” live weight of ewes in different physiological states offered varying levels of ryegrass-based pasture prior to fasting. It is important to correct for liveweight losses associated with handling and delayed weighing of sheep. The developed equations utilized recorded time by the weigh systems to adjust for weight. To use these equations if incorporated into modern weighing systems, would require manual entry of the time when ewes are removed off pasture. Providing a supplement or water during the period off pasture would likely alter the reported ewe liveweight loss patterns, probably, maintaining the “true” live weights. This might be recommended for smaller flock sizes. However, in extensive sheep production systems with an average of 2500 ewes, supplementing ewes at the weighing facility each time of weighing would have serious practical and economic implications.

5.5 Conclusion

The present study has shown that ewes lose a significant amount of live weight when feed and drinking water are restricted. The study demonstrated that the rate of ewe liveweight loss follows a predictable trajectory over a period and is influenced by herbage availability offered, pregnancy-rank and stage of pregnancy. Further, the study demonstrated in support of Chapter 4 with ewe lambs that these liveweight losses can be substantially accounted for using sets of correcting equations. These equations could be incorporated into weighing systems to quickly supply farmers accurate “without delay” ewe live weight measurements. Future studies should explore how to control the unexplained source of variation and to see if differing herbages require different equations. Further, the extent to which the live weight correcting equations can be generalized to ewes from other breeds is warranted.

Foreword to Chapters 6 to 9

Not only can live weight be affected by type of feed and feeding level (Chapters 3 to 5). It can still be affected by other confounding factors such as body size/stature fleece weight and conceptus weight. Body condition score is an alternative indicator of animal performance, which circumvents these factors. Therefore, Chapters 6 to 9 present work on the relationship between a ewe's body condition score (BCS) and live weight (LW) and other physical and physiological traits, and methodologies to predict current BCS using a ewe's live weight records. Specifically, the methodology of Chapter 6 determines the nature of association between LW and BCS at a given time point and examines the factors affecting this relationship between LW and BCS. In Chapter 7, linear regression models are deployed to predict BCS from a ewe's LW, LW-change and previous BCS record. While in Chapter 8, a set of machine learning algorithms are applied on live weight records to predict its current BCS in 43−54-month-old ewes. Chapter 9 examines if additional data (i.e. ewe wither height measurement, pregnancy status and fleece weight) in addition to LW, LW-change and previous BCS record would lead to improved BCS prediction.

Chapter 6. The effect of age, stage of the annual production cycle and pregnancy-rank on the relationship between live weight and body condition score in extensively managed Romney ewes

Chapter published as: **Semakula J, Corner- Thomas R.A, R, Morris S.T, Blair, H.T, Kenyon, P.R**. **2020**. The Effect of Age, Stage of the Annual Production Cycle and Pregnancy-Rank on the Relationship between Live weight and Body Condition Score in Extensively Managed Romney Ewes. *Animals*, 10(5):784.

Abstract

This study determined the nature of the relationship between live weight and BCS and assessed the influence of stage of the annual production cycle and pregnancy-rank on the relationship between live weight and body condition score (BCS) in Romney ewes. Data were collected from the same ewes at different ages (8−18, 19−30, 31−42, 43−54, 55−66 and ≥67 months), stages of the annual production cycle (pre-breeding, at pregnancy diagnosis, pre-lambing and weaning) and pregnancy-rank (non-pregnant, single or twin). Linear regression was determined as being sufficient to accurately describe the relationship between live weight and BCS. Across all data, a one-unit change in BCS was associated with 6.2 ± 0.05 kg live weight, however, this differed by stage of the cycle, pregnancy-rank and ewe age ($p < 0.05$). The average live weight per unit change in body condition score increased with age of the ewe, was greatest at weaning and lowest pre-lambing. Among pregnancy-ranks, the average live weight per unit change was also greater during pregnancy diagnosis than pre-lambing and was greatest among single and lowest in non-pregnant ewes. The results support the hypothesis that the relationship between live weight and BCS is affected by the interaction between stage of the annual production cycle, pregnancy-rank and ewe age.

6.1 Introduction

Body condition score (BCS) is a subjective measure which provides an estimate of an animal's soft tissue reserves, predominantly fat, and is used widely by farmers and researchers to determine the physiological state of an animal (Morris et al., 2002; Vieira et al., 2015). Body condition score was first developed for sheep (*Ovis aries*) by Jefferies (1961) and was based on a 1.0−5.0 scale, using half units. Body condition score is assessed by the palpation of the lumbar vertebrae (spinous and transverse process) immediately caudal to the last rib and above the kidneys (Jefferies, 1961; Kenyon et al., 2014). Body condition score can circumvent the shortcomings of live weight (LW), which include the effect of gut-fill, frame size, fleece weight and physiological state (Kenyon et al., 2014; Brown et al., 2015; Morel et al., 2016). Body condition score can be easily learned and is cost-effective and requires no specialized equipment (Kenyon et al., 2014). In addition, it has been suggested that BCS could be used to provide proper feeding management of a grazing flock throughout the year, detect subtle changes in condition not noticeable by visual inspection, allow farmers to be more aware of major losses in condition and be used follow changes in nutrition(Jefferies, 1961). Body condition score is thus considered a useful way for farmers to monitor the condition of their flock and estimate the required plane of nutritional allowance (Kenyon et al., 2014).

Despite the advantages of using BCS over live weight to better manage flocks, it is uncommon for producers/farmers to regularly and objectively do so. A survey of sheep producers in Australia indicated that although 96% of respondents said they monitored the body condition of their sheep, only 7% conducted hands-on BCS assessment to estimate the energy requirements of their sheep (Jones et al., 2011). In New Zealand, Corner-Thomas et al. (2016) reported that the proportion of farmers using BCS as a management tool at 40%.Combined these findings indicate that there is a sizable number of farmers not using BCS, especially in countries with large flocks. Besier and Hopkins (1989) reported that, farmers rely on a visual inspection method, that has been demonstrated to be very inaccurate or prefer to use live weight measures only. The reasons for low BCS uptake among farmers include; i) body condition score being subjective, depending on the judgement of the assessor; 2) it is labour intensive and 3) requires training of the assessors, who should regularly undergo recalibration (Kenyon et al., 2014). Strategies to increase the adoption and use of BCS among farmers and the reliability of measures included; promotional farmers' training workshops and regular assessor recalibration (Kenyon et al., 2014). However, given the apparent low rate of farmer use, these strategies appear not to have yielded the desired outcome presumably because they do not directly address how to reduce the labour burden associated with hands-on BCS. Therefore, it is argued that, consistent and accurate alternative methods to estimate body condition score of sheep that require less hands-on measurement would likely be advantageous and improve uptake and use. Ideally, this prediction would be based on a management tool already utilized on farm, so that it reduces workload, it would be quick and not subjective in nature.

Body condition score is correlated with live weight and have been reported to have either a positive linear relationship (Kenyon et al., 2014; Morel et al., 2016) or a curvilinear relationship in ewes (Teixeira et al., 1989). Factors such as breed, frame size, composition and patterns of fat distribution in the body (Kenyon et al., 2014; McHugh et al., 2019) have been reported to affect the average change in live weight associated with a one-unit change in BCS. The magnitude of the relationship between BCS and liveweight changes and with physiological status, age and breed of sheep (Sezenler et al., 2011; McHugh et al., 2019). Data on changes in either live weight or BCS reflects changes in an animal's body condition and can be used to inform decisions on appropriate feed allocation at a given physiological status and breed (Keady et al., 2005). Therefore, assessment of the relationship between live weight and BCS can be a valuable tool to maximize animal productivity and feed utilization (Roche et al., 2006; Morel et al., 2016). The relationship between live weight and BCS has generally been described by simple linear regression (based on R^2) likely due to the simplest linear relationship appearing to be as strong as more complex models. However, using the coefficient of determination of a regression alone, as the criterion for goodness-of-fit, is not suitable to validate models because it does not provide information about the degree to which the predicted values diverge from true values (Goopy et al., 2018; Wamatu et al., 2019). Moreover, models should be robust in predicting other datasets. The majority of the previous studies have been based on fixed BCS ranges (mostly from 2.5 to 4.0) and it is unclear whether such a strong relationship would be observed in a wider range of BCS values (1 to 5). To date no known attempts have been made to establish the true nature of the relationship between LW and BCS using a whole range of BCS values. It was hypothesized that the relationship between LW and BCS would be adequately described by a linear regression.

In cattle, the average liveweight change (gained or lost weight) associated with each BCS one-unit change is well associated with BCS (Berry et al., 2007; McHugh et al., 2019). Similar adjustment factors for sheep, however, have received less investigation. The current BCS adjustment factors in sheep have been generated from either relatively small-scale studies (n = 28, Morel et al., 2016; n = 156, Sezenler et al. 2011) or single time point observations (point specific) based on within-flock studies (Sezenler et al., 2011; Kenyon et al., 2014; McHugh et al., 2019). Ideally, the relationship between BCS and live weight should be investigated using the same individuals over time. To these authors knowledge no studies have conducted longitudinal studies for this purpose. Both conventional and modern weighing systems combined with individual electronic identification can now allow lifetime data to be collected more easily and quickly on large sheep flocks. Using this technology combined with an individual BCS at a given point in their lifetime, therefore, can allow specific stage of life BCS live weight relationship to be developed. It was hypothesized that the relationship between live weight and BCS would be modified by stage of annual production cycle and pregnancy-rank over time. Therefore, this study had three objectives: i) to determine the nature of the relationship between live weight and BCS, using both coefficient of determination and prediction error; ii) to quantify the average liveweight change associated with each incremental change in BCS on a scale from 1.0 to 5.0, with 0.5-point intervals; and iii) to determine if the association differed by stage of the annual production cycle, pregnancy-rank, and over time in Romney ewes.

6.2 Materials and Methods

6.2.1 Farms and animals

The current study utilized datasets from a database collected between 2011 and 2015. Data were collected as part of normal routine farm management from two commercial New Zealand sheep farms in which all ewes were bred as ewe-lambs at approximately eight months of age at breeding. Farm A was located in the Waikato region of New Zealand and consisted of Romney ewes. Two cohorts of ewes from Farm (A) were included in this study: 2010-born (n = 3469) and 2011-born (n = 4572). Farm (B) was located in the Wairarapa region of New Zealand, with Romney ewes that were born in 2011 (n = 3760). The number of ewes monitored on each farm fluctuated by stage of the annual cycle. This was influenced by each farm owner's decision to keep or dispose (cull) of ewes or failure to collect data during any period. Farm (A) did not collect live weight and BCS data during the pre-lambing period on two occasions.

All ewes were weighed to the nearest 0.2 kg using static digital weighing scales (Tru-Test group, model XR5000) and were body condition scored (BCS) at the same time. Body condition score was assessed by palpating the soft tissue over the lumbar region on a 1.0−5.0 scale (1 = emaciated, 5 = obese) assessed to the nearest 0.5 unit (Jefferies, 1961; Kenyon et al., 2014). BCS was assessed immediately prior to breeding (two to three weeks before start of mating), at pregnancy diagnosis (approximately 80 days after start of mating), pre-lambing (within three weeks before start of lambing) and at weaning (approximately 100 days after start of lambing). Body condition was measured over 6 years, beginning at approximately 8 months of age (age groups: 8−18, 19−30, 31−42, 43−54, 55−66, ≥67 months). Body condition score was determined by two experienced assessors (one for first 6 years (2011−2016) and one for the final year

(2016)). The timing of measurements and the number of animals measured are summarized in Tables 6.1 and 6.2. Additional information collected included farm, year of observation, pregnancy-rank and age. The pregnancy-rank of the ewes was determined using transabdominal ultrasound conducted by a commercial operator (non-pregnant (0), single foetus (1), twin (2)).

Table 6.1 Number of ewes by age group (months), stage of the annual cycle (pre-breeding, pregnancy diagnosis, pre-lambing, weaning) and farm (A, B).

Stage of the annual cycle	Age group (months)	Farm A	Farm B	Overall
Pre-breeding	8–18	8046	3752	11798
	19–30	5110	3626	8736
	31–42	3884	3027	6911
	43–54	3043	2294	5337
	55–66	2504	1921	4425
	≥67	444	1044	1488
Pregnancy diagnosis	8–18	7635	3760	11395
	19–30	4805	3489	8294
	31–42	3607	2961	6568
	43–54	2882	2241	5123
	55–66	2185	1829	4014
	≥67	477	919	1396
Pre-lambing	8–18	6508	1624	8132
	19–30	2382	3225	5607
	31–42	NA	2840	2840
	43–54	1461	1867	3328
	55–66	1034	1759	2793
	≥67	NA	930	930
Weaning	8–18	5039	3708	8747
	19–30	4062	3177	7239
	31–42	3100	2661	5761
	43–54	2580	1986	4566
	55–66	1658	1112	2770
	≥67	33	564	597

NA: data not collected

Table 6.2 Number of ewes by pregnancy-rank (Non-pregnant, Single foetus, Twin), age group (months) and farm (A, B) during the different stages of the annual cycle (Pregnancy diagnosis, Pre-lambing).

		Pregnancy diagnosis		Pre-lambing		
Pregnancy-rank	Age group (months)	Farm A	Farm B	Farm A	Farm B	Overall
Non-pregnant	8−18	1051	482	NA	NA	1533
Non-pregnant	19−30	120	229	NA	NA	349
Non-pregnant	31−42	55	70	NA	NA	125
Non-pregnant	43−54	40	95	NA	NA	135
Non-pregnant	55−66	78	50	NA	NA	128
Non-pregnant	≥ 67	68		NA	NA	68
Single	8−18	3277	978	3229	957	8441
Single	19−30	1287	1952	571	1890	5700
Single	31−42	1038	1363	NA	1348	3749
Single	43−54	650	854	267	798	2569
Single	55−66	324	767	258	755	2104
Single	≥ 67	83	204	NA	181	468
Twin	8−18	3310	652	3249	637	7848
Twin	19−30	3400	1315	1803	1262	7780
Twin	31−42	2501	1535	NA	1498	5534
Twin	43−54	2185	1299	1185	1065	5734
Twin	55−66	1765	1019	768	981	4533
Twin	≥ 67	284	722	NA	692	1698

NA: data not collected

6.2.2 Data Management

Live weight and BCS data were first exported to Microsoft excel version 2010 for pre-processing including cleaning, merging and validation. Data were then exported to the R statistical program version 3.3.4 (R Core Team, 2016) for further management. A total of 128,753 records from 11,798 ewes were collected between 2011 and 2016 (Tables 6.1 and 6.2). Records were removed from the analysis dataset that had no identification (n = 15) or that had live weight for the calibration weights (test weights) recorded (n = 9), so were removed from the analysis. The independent variables included: age group, determined by number of months at the time of breeding time within a 12 month period (i.e. 8−18, 19−30, 31−42, 43−54, 55−66 and ≥67 months); stage of annual production cycle (pre-breeding, at pregnancy diagnosis, pre-lambing and weaning); and pregnancy-rank (non-pregnant: 0, single foetus: 1, twin: 2). In both farms, triplets (n = 67) were not considered due to fewer numbers and high variability in both live weight and BCS compared with their contemporaries. A variable labelled FarmYear was generated to account for the different birth years as well as farm of origin. When live weight was considered as the dependent variable, BCS was considered its covariate and vice versa.

6.2.3 Statistical analyses

All analyses were conducted in R program version 3.4.4 (R Core Team, 2016). Pearson's correlation between BCS and live weight was estimated across all data and within each age group and stage of the annual production cycle. Correlation coefficients were also estimated for each age group with adjustment for stage of the annual production cycle and pregnancy-rank (for measurements made at pregnancy diagnosis and pre-lambing). Any significant differences between correlation coefficients were determined based on Fisher's r-to-z transformation.

6.2.4 Model development and selection.

To determine the true nature of the relationship between live weight and BCS, linear (LM), second order polynomial/quadratic (QUAD), Box-Cox and square root (SQRT) transformation regressions were compared. The best lambda (with greatest likelihood) for Box-Cox transformation was 0.67. Table 6.3 gives the formulae by which the models, their goodness of fit or coefficient of determination (r^2: for simple and R^2: for multiple regression) and error metric (Mean Absolute Error and Mean Absolute Percent Error) were defined (Moriasi et al., 2007; Li, 2017; Botchkarev, 2019). For this comparison, the percent error and the goodness-of-fit were based on the testing dataset. The models were adjusted for the effects of stage of the annual production cycle, age group and FarmYear. The models were examined for normality of the residuals and heteroscedasticity and outliers were examined using residual plots. In addition, Cooks distances were calculated for each model to assess the existence of outliers that may have influenced coefficients of the models. The leverage plots were used to detect data points with unusually high influence (Cook, 1977). Outliers highlighted on the diagnostic plots were investigated and corrected if identified as a simple typing error or removed. The resulting dataset was then reanalyzed to determine its influence. In total, six of the 128,753 live weight data points were removed.

Table 6.3 Formulae for live weight estimation models (Linear model (SLM), Quadratic transformation (QUAD), Box_Cox transformation (Box_Cox) and Square root transformation (SQRT)) using body condition score (BCS), adjusted R^2, error metrics (Mean Absolute Error: MAE; Percent Error: PE) and coefficient of variation of the live weight (LW).

Model / measure	Formula
Linear model (SLM)	$\mathrm{LW} = \alpha + \mathrm{BCS}$
Quadratic transformation (QUAD)	$\mathrm{LW} = \alpha + \mathrm{b(BCS)} + \mathrm{c(BCS)}^2$
Square root transformation (SQRT)	$\mathrm{LW}^{0.50} = \alpha + \mathrm{BCS}$
Box_Cox transformation (Box_Cox)	$\mathrm{LW}^{0.67} = \alpha + \mathrm{BCS}$
Coefficient of determination (R^2)	$\mathrm{R}^2 = 1 - \frac{\mathrm{MS_{RES}}}{\mathrm{MS_{Tot}}}$
Adjusted R^2 (Adj.R^2)	$\mathrm{Adj.R}^2 = 1 - [\frac{(1-\mathrm{R}^2)+(\mathrm{n}-1)}{\mathrm{n}-\mathrm{k}-1}]$
Mean Absolute Error (MAE)	$\mathrm{MAE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{n}}(\lvert \mathrm{y_j} - \hat{\mathrm{y}}_\mathrm{j} \rvert)$
Mean Absolute Error Percent (PE)	$\mathrm{PE} = \frac{1}{\mathrm{n}}\sum_{\mathrm{j}=1}^{\mathrm{n}}\left(\left\lvert \frac{\mathrm{y_j}-\hat{\mathrm{y}}_\mathrm{j}}{\mathrm{y_j}} \right\rvert\right) * 100$
Coefficient of variation (CV)	$\mathrm{CV} = 100\sqrt{\frac{\mathrm{MS_R}}{\mathrm{mean}}}$

α indicates the intercept. *bc* indicate the regression coefficients. y_j indicates the actual expected output. $\hat{y}_j$ indicates the model's prediction. *k* indicates the number of independent predictors. *n* indicates the sample size/number of data points. MS_R indicates variation due to the model. MS_{Tot} indicates total variation.

6.2.5 Final model fitting (factors affecting the relationship between LW and BCS)

The best linear model for final data fitting was selected by comparing two parameter estimation methods (a generalized least squares vs linear mixed-effects model). The linear mixed effects model (LMM) was selected for fitting the model, as it had the smallest likelihood value and Akaike's Information Criterion (AIC) values (p < 0.001). To quantify the relationship between live weight and BCS and the factors associated with this relationship, the final analysis was based on the minimal LMM model (with minimum Akaike's Information Criterion, AIC value retained during simplification) incorporating all significant effects using the nlme package (Pinheiro et al., 2018). Three separate live weight estimating models were constructed. The first model included body condition score (BCS) as a covariate, age group (A) and stage of the annual production cycle (T) as explanatory variables. To determine the impact of pregnancy-rank, two additional models (one for measurements at pregnancy diagnosis and another pre-lambing) were constructed, each of the models taking a similar form. In both models, BCS was treated as a covariate, age group and pregnancy-rank (P) as explanatory variables. To test whether BCS effects on live weight were modified by age group, stage of the annual production cycle and pregnancy-rank, the models included up to three-way interactions (BCS x A x T or BCS x A x P). A similar approach was used when assessing the effect of all other factors on BCS. FarmYear and individual ewe electronic identification number (EID) were included as random variables. Variance functions to account for heteroscedasticity and an auto regressive temporal correlation structure to account for temporal dependency of nearby stage of the annual production cycle were also included. The differences among intercepts and slopes (beta coefficients) in the model were compared using Tukey's pairwise contrasts on the final model using the multcomp, package (Hothorn et al., 2008). Statistical significance from the model using ANOVA type III are reported. To estimate the least squares means for BCS, the models above were refitted using BCS as the dependent variable and LW considered its covariate.

All models were constructed, fitted and cross-validated using machine learning algorithms, implemented in four steps. The steps included i) data partitioning, ii) resampling, iii) model training and iv) validation. Data partitioning involved dividing the initial dataset (with stratification preserving the class proportions) into training and testing datasets in a ratio of 3:1, with replacement. Resampling involved using bootstrapping and aggregation procedures (Breiman, 1996; Tropsha et al., 2003) to select 10 subsamples from the training set and repeating the resampling five times. Model training involved fitting of the linear regression using the training dataset subsamples (10) from which, nine were used for computing the parameters (i.e. β) while the remaining one part was used for error estimation (ε). Finally, all parameters were to determine the final value (estimate). Model cross-validation involved using the trained model to predict BCS in the testing dataset.

6.3 Results

A total of 128,753 ewe records were included in the analysis (Tables 6.1, 6.2). The number of records (n) decreased with ewe age. The majority of ewes were diagnosed as pregnant (93.3%, n = 32,764) with more ewes carrying twin foetus (56.9%, n = 19,987) compared with single (36.9%, n = 12,777) (Table A 6.3). Body condition scores of 3.0 (41.6%, n = 56,381) and 2.5 (39.4%, n = 53,470) formed the bulk of the records while 1.0 (0.0%, n = 19) and 5.0 (0.0%, n = 6) were the least frequent (Table 6.4). The overall mean live weight of ewes in this study was 54.2 kg (SD = 9.3 kg) and BCS was 2.81 (SD = 0.42). There was relatively high variability in live weight for each BCS (mean CV = 15%, Table 6.4).

Table 6.4 Table 6.4 Number of records (n), mean and standard deviation (SD), coefficient of variation (CV %) for live weight across BCS.

BCS (units)	Number of records (n)	Live weight (kg) Mean	SD	CV %
1.0	19	41.5	5.6	13.4
1.5	350	45.6	8.1	17.7
2.0	7735	49	8.1	16.6
2.5	53470	51.6	8.6	16.7
3.0	56381	55.8	8.8	15.7
3.5	15051	59.4	9.7	16.4
4.0	2350	62.2	10.7	17.2
4.5	241	60.6	11.2	18.5
5.0	6	67.8	3.6	5.3

6.3.1 Nature of association between live weight and BCS

The models were more stable at BCS from 2.5 to 3.5 (i.e. all model lines of best fit converged, Figure 6.1). All models had comparable statistical parameters (μ, SD) and not significantly different from the observed data (Appendix VII Figure 1a). Examination of the diagnostic plots for all four models that assessed the nature of the relationship between live weight and BCS (in the initial construction) revealed that at the tails of the datasets were "hanging" (not lying on the diagonal QQplot line, Appendix VII Figure 1b) an indication that all the models were sensitive to bias at the extremes of the dataset. In addition, all models had relatively similar goodness-of-fit (R^2 = 0.69) and Cook's distances suggesting relatively similar robustness of these models to outlier effects (Table 6.5).

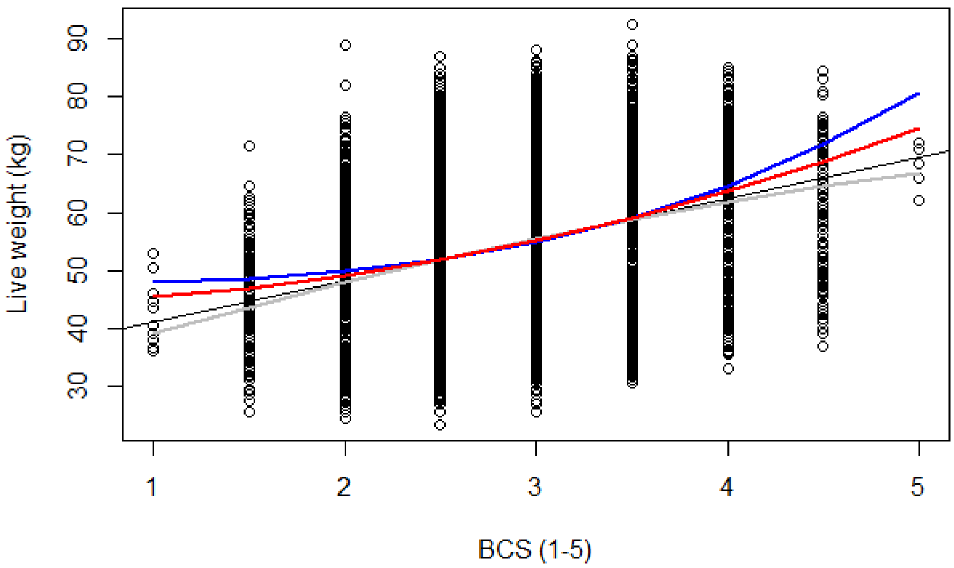


Figure 6.1 Ewe live weight (kg) as a function of BCS (1.0−5.0). Line of best fit is given for (linear model (SLM): black colour, Quadratic transformation (QUAD): grey, Box-Cox transformation (Box_Cox): red and square root transformation (SQRT): blue)

Table 6.5 Mean Absolute Error (MAE) , Percentiles of Percentage error, Adjusted R^2 and percentiles of Cook's distance of the models (Linear model (SLM), Quadratic transformation (QUAD), Box-Cox transformation (Box_Cox) and Square root transformation (SQRT)) for live weight predictions on testing dataset.

Model	SLM	QUAD	SQRT	Box_Cox
Adjusted R^2	0.69	0.69	0.69	0.69
MAE	4.12	4.11	0.28	0.74
P-value	***	***	***	***
Percentiles of PE				
75th	7.8%	7.7%	7.8%	7.8%
90th	7.8%	7.8%	7.8%	7.8%
95th	7.8%	7.8%	7.8%	7.9%
Percentiles of Cook's distance				
75th	0.00001	0.00001	0.00001	0.00001
90th	0.00003	0.00002	0.00003	0.00003
95th	0.00005	0.00004	0.00004	0.00004

*** indicate significance at $p < 0.001$

6.3.2 Effect of age, stage of annual production cycle and pregnancy-rank on ewe LW and BCS

Age group, stage of annual production cycle and pregnancy-rank all affected ewe LW ($p < 0.05$). As ewes increased in age their live weight increased ($p < 0.05$) across all stages of the annual production cycle, plateauing after 55−66 months (Appendix VII Figure 2a). Ewes were heaviest ($p < 0.01$) at pre-lambing in their last year of observation (≥67 months). Within age (except at 8−18 months), ewes were consistently heaviest ($p < 0.05$) at pre-lambing. Among pregnancy-ranks, live weight of ewes varied differently over time ($p < 0.05$) with no clear pattern observed (Appendix VII Figure 2b). There was, however, more variability in the live weights of non-pregnant ewes than those bearing singles or twins. At pregnancy diagnosis, live weight was lowest ($p < 0.05$) in non-pregnant ewes in the first four age groups (8−18, 19−30, 31−42 and 43−54 months) compared with their contemporaries. Twin-bearing ewes consistently had greater ($p < 0.05$) live weight than single or non-pregnant ewes across age up to the 43−54 months. Pre-lambing, live weight was greater in twin than single-bearing ewes ($p < 0.01$) up to the 55−66 months.

Body condition score was influenced ($p < 0.01$) by age, stage of annual production cycle and pregnancy-rank ($p > 0.05$). Body condition score decreased as the ewe increased in age ($p < 0.05$) across all stages of the annual production cycle plateauing after 55−66 months (Appendix VII Figure 3a). However, when disaggregated by stage of the animal cycle and pregnancy-rank, BCS tended to decrease among the non-pregnant ewes at pregnancy diagnosis but with no clear pattern among other ranks (Appendix VII Figure 3b). With the exception of age groups 31−42 and 55−66 months, across age, BCS was lowest ($p < 0.05$) pre-lambing. Within the annual production cycle and over time, the BCS of ewes showed no clear pattern of decline. Among pregnancy-ranks, BCS at pregnancy diagnosis was greater ($p < 0.05$) in the first two years (8−18, 19−30 months) after which, it decreased remaining comparable ($p > 0.05$) among pregnancy-ranks. Pre-lambing, BCS was greater in single than twin-bearing ewes across age except for ≥67 months.

6.3.3 Effect of age, stage of annual production cycle and pregnancy-rank on the relationship between live weight and BCS

Overall, the correlation between BCS and live weight was 0.47, indicating that 21% ($R^2 = 0.21$) of the variability in live weight was explained by differences in BCS. When adjusted for age, stage of annual production cycle and pregnancy-rank, the overall correlation decreased slightly to 0.44 ($R^2 = 0.18$). The correlation between BCS and live weight was affected by both age of the ewe, stage of the annual production cycle and pregnancy-rank ($p < 0.05$). Overall, the correlation between live weight and BCS varied from 0.02 pre-lambing to 0.69 at pregnancy diagnosis (Tables 6.6, 7).

The strength of the association between BCS and live weight differed significantly ($p < 0.05$) across both age of ewe and stage of the annual production cycle. Within age group, the correlation between live weight and BCS was relatively similar except at ≥67 months. Within stage of the cycle, the correlation between live weight and BCS was strongest at weaning and weakest at pre-lambing. Within pregnancy-rank, the correlation between live weight and BCS varied from 0.02 ($p > 0.05$) pre-lambing to 0.69 ($p < 0.01$) at pregnancy diagnosis. There was no clear pattern in the strength of association among age groups and pregnancy-ranks ($p > 0.05$).

Tables 6.6 and 6.7 summarize the regression equations of the relationship between live weight and BCS by age of ewe, stage of the cycle and pregnancy-rank. The regression intercepts, as well as the average change in live weight per one-unit change in BCS (incremental liveweight change), were affected by all three factors ($p < 0.05$).

The incremental liveweight change increased ($p < 0.001$) as the ewes aged. The magnitude of the incremental liveweight change of ewes in the same age group (Table 6.6) was altered by the stage of the annual production cycle ($p < 0.001$). The increase in the average incremental liveweight change varied from 2.3 kg for younger ewes pre-lambing (8−18 month) to 9.5 kg for the older ewes at weaning (≥67 month).

Within stage of the annual production cycle, the incremental liveweight change was lowest ($p < 0.01$) at 8−18 month but the maximum change varied by stage of the annual production cycle for example at ≥67 months for pre-breeding and at pregnancy diagnosis, 43−54 for pre-lambing and 55−66 for weaning. Weaning was associated with the greatest incremental liveweight change (5.6 to 9.5 kg) while pre-lambing was associated with the lowest (2.3 to 5.9 kg) (p< 0.05).

Table 6.6 Intercepts (α), coefficients (β), correlation coefficient (r_{xy}) and adjusted R^2 for the regression of the live weight with body condition score for each stage of the annual production cycle (pre-breeding, at pregnancy diagnosis, pre-lambing, weaning) and ewe age (8−18 months, 19−30, 31−42, 43−54, 55−66 and ≥67).

Stage of annual production cycle	Age group	α(*SE*)	β(*SE*)	r_{xy}	Adj.R^2
Pre-breeding	8−18	33.2(0.25)	2.8(0.09) ^{a}	0.43^{bc}	0.15
	19−30	36.5(0.29)	6.0(0.10) ^{d}	0.49^{bc}	0.24
	31−42	36.9(0.36)	7.1(0.13) ^{ef}	0.50^{bc}	0.26
	43−54	39.5(0.4)	6.9(0.13) ^{e}	0.48^{bc}	0.28
	55−66	46.0(0.39)	5.8(0.14) ^{d}	0.48^{bc}	0.23
	≥67	37.6(0.72)	8.4(0.23) ^{g}	0.58^{c}	0.35
At pregnancy diagnosis	8−18	34.9(0.25)	2.8(0.09) ^{a}	0.41^{bc}	0.13
	19−30	35.6(0.32)	5.0(0.12) ^{c}	0.34^{b}	0.15
	31−42	38.3(0.35)	5.9(0.12) ^{d}	0.49^{bc}	0.26
	43−54	38.4(0.41)	7.0(0.14) ^{ef}	0.45^{bc}	0.21
	55−66	40.8(0.49)	7.0(0.17) ^{ef}	0.45^{bc}	0.23
	≥67	42.1(0.72)	7.2(0.22) ^{ef}	0.56^{c}	0.31
Pre-lambing	8−18	42.6(0.34)	2.3(0.12) ^{a}	0.06^{a}	0.24
	19−30	50.5(0.38)	2.4(0.14) ^{a}	0.14^{a})	0.06
	31−42	48.9(0.54)	4.0(0.19) ^{b}	0.29^{a})	0.1
	43−54	48.3(0.44)	5.9(0.16) ^{d}	0.13^{a}	0.21
	55−66	52.2(0.61)	5.3(0.21) ^{cd}	0.13^{a}	0.1
	≥67	57.2(0.92)	4.8(0.35) ^{bcd}	0.32^{ab}	0.07
Weaning	8−18	30.9(0.25)	7.5(0.09) ^{f}	0.57^{c}	0.45
	19−30	38.3(0.27)	5.6(0.09) ^{d}	0.57^{c}	0.28
	31−42	35.9(0.34)	7.4(0.11) ^{ef}	0.58^{c}	0.36
	43−54	36.1(0.38)	8.3(0.14) ^{g}	0.62^{c}	0.3
	55−66	34.8(0.43)	9.5(0.16) ^{h}	0.62^{c}	0.4
	≥67	39.8(0.86)	7.5(0.3) ^{efg}	0.64^{cd}	0.41

$^{a-n}$, superscripts within column indicate significant difference at p < 0.05. *SE* denotes standard error

Among pregnancy-ranks, the increase in incremental liveweight change was greater (p < 0.01) at pregnancy diagnosis (4.3 to 13.6 kg) compared with pre-lambing and increased with age of ewe. Pre-lambing, the increase in incremental liveweight change had no clear pattern. Generally, incremental liveweight change for similar age groups appears to have varied randomly regardless of pregnancy-rank (Table 6.7). At pregnancy diagnosis, the incremental liveweight change was greater (p < 0.05) in single- and twin-bearing ewes than non-pregnant ewes at all age groups except at 8−18 and ≥67 months. The incremental liveweight change was also comparable (p > 0.05) for both single- and twin-bearing ewes except at 19−30 months. Pre-lambing, the incremental liveweight change was unexpectedly low (0.4 to 3.8 kg) and varied with no clear pattern among pregnancy-ranks as the ewe aged.

Table 6.7 Intercepts (α), coefficients (β), correlation coefficient (r_{xy}) and adjusted R^2 for the regression of the live weight with body condition score value for each age (8−18 months, 19−30, 31−42, 43−54, 55−66 and ≥67) by pregnancy-rank (non-pregnant, single and twin bearer) and stage of the annual production cycle (at pregnancy diagnosis and pre-lambing).

Pregnancy-rank	Age group (months)	α(se)	β(se)	r_{xy}	Adj. R^2
At pregnancy diagnosis					
Non-pregnant	8−18	21.0(1.17)	9.4(0.41) ^{d}	0.59^{c}	0.06
Single		29.9(0.54)	7.5(0.19) ^{b}	0.37^{b}	0.05
Twin		29.3(0.55)	7.9(0.19^{b}	0.44^{ab}	0.17
Non-pregnant	19−30	23.4(2.43)	8.3(0.9) ^{de}	0.69^{cde}	0.15
Single		27.2(0.56)	12.1(0.2) ^{h}	0.56c	0.16
Twin		31.6(0.47)	7.2(0.16) ^{b}	0.41^{b}	0.12
Non-pregnant	31−42	36.7(3.08)	4.3(1.14) ^{a}	0.38^{c}	0.43
Single		20.9(0.62)	10.9(0.29) ^{ef}	0.53^{bc}	0.31
Twin		26.8(0.48)	8.9(0.17) ^{cd}	0.47^{bc}	0.28
Non-pregnant	43−54	28.8(2.43)	8.0(0.87) ^{c}	0.36^{b}	0.47
Single		20.5(0.75)	11.1(0.26) ^{ef}	0.52^{b}	0.2
Twin		23.7(0.51)	10.1(0.18) ^{ef}	0.48^{b}	0.16
Non-pregnant	55−66	27.5(2.51)	8.7(0.88) ^{ef}	0.53^{b}	0.17
Single		19.1(0.84)	11.6(0.29) ^{fg}	0.50^{ab}	0.15
Twin		22.8(0.55)	10.4(0.19) ^{e}	0.49^{b}	0.15
Non-pregnant	≥ 67	30.4(2.31)	7.8(0.80) ^{bc}	0.32^{cd}	0.27
Single		18.4(1.61)	11.8(0.55) ^{g}	0.47^{c}	0.34
Twin		13.2(0.85)	13.6(0.29) ^{hi}	0.52^{c}	0.27
Pre-lambing					
Non-pregnant	8−18				
Single		52.0(0.73)	1.2(0.26) ^{c}	0.06^{a}	0.02
Twin		45.9(0.78)	3.8(0.29) ^{e}	0.04^{a}	0.01
Non-pregnant	19−30				
Single		51.6(0.74)	1.6(0.46) ^{cd}	0.05^{b}	0.01
Twin		52.7(0.63)	1.3(0.43) ^{cd}	0.06^{b}	0.12
Non-pregnant	31−42				
Single		52.2(0.8)	1.6(0.29) ^{cd}	0.04^{bc}	0.02
Twin		52.7(0.64)	1.5(0.23) ^{cd}	0.06^{b}	0.03
Non-pregnant	43−54				
Single		50.5(0.96)	2.2(0.35) ^{cd}	0.11^{ab}	0.01
Twin		51.7(0.66)	2.0(0.24) ^{cd}	0.02^{b}	0.1
Non-pregnant	55−66				
Single		54.4(1.02)	0.9(0.36) ^{abc}	0.06^{ab}	0.08
Twin		50.1(0.71)	2.6(0.26) ^{de}	0.05^{ab}	0.04
Non-pregnant	≥ 67				
Single		50.3(1.94)	2.4(0.69) ^{cde}	0.15^{c}	0.02
Twin		58.8(1.03)	0.4(0.37) ^{ab}	0.02^{b}	0.01

$^{a\text{-}j}$ Different superscripts within column and stage of annual production cycle indicates differences at $p < 0.05$.

6.4 Discussion

This study was aimed to determine the nature of the association between live weight and BCS and to quantify the average liveweight change associated with each incremental change in ewe BCS as measured on a 1.0 to 5.0 scale with 0.5-point intervals. In addition, the extent to which this association differed by stage of the annual production cycle and pregnancy-rank and ewe age, in extensively managed Romney ewes was investigated. It was hypothesized that the relationship between BCS and live weight was best described using a linear regression and would vary based on age group, stage of cycle and pregnancy-rank.

In the present study, the linear regression was considered sufficient to describe the relationship between live weight and body condition score. This was not surprising as the majority of previous studies have reported a linear relationship between live weight and BCS (Kenyon et al., 2014). In addition, transforming data would add unneeded complexity to the model (Lazar, 2010). The percent error for all the four models (LM, QUAD, Box-Cox and SQRT) was within acceptable range (i.e. < 10%), for veterinary purposes (Leach and Roberts, 1981) and prediction models (Alexander et al., 2015). The findings show, therefore, that live weight and BCS vary together in a linear manner and this relationship ship can be predictable using simple linear regression.

Live weight increased with the ewe age and began to plateau at 43 months of age. Ewe live weight increases with frame size when as an animal ages, until its mature size is achieved (Ho et al., 1989). In temperate (European) sheep breeds, this has been reported to occur between 25−50 months of age (Wiener, 1967; Cake et al., 2006). The present findings are in agreement with other authors who reported a live weight increase with age, plateauing after 33 months of age in Romney ewes (Loureiro et al., 2012; Pettigrew et al., 2019). In three thin-tailed breeds of indigenous Turkish sheep live weight increased with age (Sezenler et al., 2011), although that study did not have age groups below two years to demonstrate the overall trend.

Within age group, ewe live weight was highest at pre-lambing. During late pregnancy, the conceptus weight influences total ewe live weight, as single, twin and triplets near term can add 5−8 kg, 12−17 kg and 17−21 kg, respectively (Kenyon et al., 2007b; Loureiro et al., 2010). Thus, it was perhaps not surprising that ewe live weight was heaviest pre-lambing and tended to be increase with pregnancy-rank. The observed low live weight among the non-pregnant ewes in the first three years, particularly during at pregnancy diagnosis may be explained by the fact that lighter ewes are less likely to conceive.

There was a general decline in BCS with age of the ewe, which began to plateau from 55 months of age across the stages of the annual production cycle. This finding is in agreement with a declining trend for BCS with age at breeding in Merino and Corriedale ewes (Gonzalez et al., 1997). However, it has been reported that thin-tailed breeds of indigenous Turkish sheep had greater BCS scores pre-breeding but lower scores at lambing and weaning across age groups (Sezenler et al., 2011). The results of the current study contrast with others who reported greater condition scores as a ewe aged across all stages of the annual production cycle in mature mixed sheep breeds and crossbreds (McHugh et al., 2019). Breed differences and nutritional conditions may explain the differences observed between studies. However, we did not collect data on nutritional status due to the extensive nature of the study and given that was conducted over multiple seasons and years. A declining BCS over time indicates that ewes used their body reserves to meet their nutritional demands, thus, suggesting that at times ewes in this study were likely not being fed to meet their theoretical nutritional requirements, particularly in lactation. The change in trend of BCS when data were disaggregated by pregnancy-rank highlights a potential interaction between factors affecting BCS in sheep. The ewes found to be non-pregnant during pregnancy diagnosis in the first two age groups (8−18 and 19−30 months) were also lighter. The finding therefore agrees with previous studies (Kenyon et al., 2004a, 2004b; Corner-Thomas et al., 2015a).

The correlation between BCS and live weight was weak to moderate based on scale of 0 to 1.0 (Chan, 2003; Akoglu, 2018), ranging between 0.18 and 0.67 across ages, stages of the annual production cycle and pregnancy-rank. By adjusting for age and stage of the annual production cycle, these results suggest that 6% to 45% of the variability in live weight was explained by differences in BCS and vice versa. These values are lower than those previously reported by others 0.60 to 0.82 (Sezenler et al., 2011) from data of 156 ewes and 0.81 from data of 28 mixed aged Romney ewes (Morel et al., 2016). They are comparable, however, to those reported for a study with multiple breeds (0.36 to 0.63) and stages of the annual production cycle (0.42 to 0.62) (McHugh et al., 2019). The between studies difference in correlation strength may be explained by variation in sample sizes, breed, stage of the annual production cycle and study design. The weaker correlation between BCS and live weight observed at pre-lambing could be attributed to the difficulty (data can be more variable) to body condition score heavily pregnant ewes (Yates and Gleeson, 1975; Kenyon et al., 2014).

In this study, a linear relationship between live weight and body condition score was demonstrated. This relationship was affected by ewe age, stage of the annual production cycle and pregnancy-rank. These results are in agreement with previous findings showing significant age and stage of the annual production cycle effects (Kenyon et al., 2014; Morel et al., 2016; McHugh et al., 2019). A linear relationship suggests that, for a given breed type, a single incremental liveweight change across the entire BCS range can be applied. The incremental liveweight change increased with age of ewe and varied across stage of the annual production cycle being numerically lowest at 8−18 months and greatest at ≥ 67 months. Thus, as a ewe ages, a greater liveweight change is required to alter BCS by one unit, which translates into greater energy requirements in order to make the change (Freer et al., 2007; Morel et al., 2016).

The relationship between BCS and live weight also varied by stage of the annual production cycle. Overall, the liveweight change required to cause a one-unit change in BCS was greatest at weaning and lowest at pre-lambing. It is not clear why the regression coefficients of live weight on BCS at pre-lambing were consistently low. It may have been because the conceptus and uterine mass was not accounted for which is likely to have confounded the true liveweight change associated with a unit change in BCS. The conceptus mass has an influence on total ewe live weight from mid-pregnancy (Kenyon et al., 2008; Kenyon et al., 2011b) which coincides with pregnancy diagnosis. Additionally, it may have been due to the difficulty associated with body condition scoring of pregnant animals (Yates and Gleeson, 1975; Kenyon et al., 2014) as previously stated. Among mature ewes (≥43 months), the incremental liveweight change during mating/breeding was within the range reported for mixed-age Romney ewes (Kenyon et al., 2014; Morel et al., 2016), but were greater than Romney composite ewes (Kenyon et al., 2004a, 2004b; Kenyon et al., 2014).

Pregnancy-rank significantly affected the live weight of ewes, their body condition scores (BCS) and eventually the relationship between live weight and BCS. The effect of pregnancy-rank on live weight was not surprising given that ewe live weight was potentially confounded by conceptus weight from mid to late pregnancy. The effect of pregnancy-rank on BCS is in agreement with earlier findings in Romney sheep (Kenyon et al., 2004b), merino sheep (Kleemann and Walker, 2005), Cheviots (Gunn et al., 1988; Gunn et al., 1991), and in Scottish blackface ewes (Rhind et al., 1984a).

The finding that the incremental liveweight change was lower in non-pregnant ewes at pregnant diagnosis was not surprising as their energy demand would be expected to be lower than for pregnant ewes. The energy demand is greater for pregnant ewes and increases with the number of foetuses (Nicol and Brookes, 2007). It is, however, not clear why the incremental liveweight change at pre-lambing varied randomly. The unexpectedly low incremental liveweight change among pregnancy-ranks at pre-lambing could have resulted from the confounding effect of the fully-grown conceptus weight.

6.5 Conclusion

In conclusion, the current study demonstrated that in a large population of ewes across a full range of BCS, live weight and BCS were linearly related and the relationship depended on the age of ewe, stage of the annual production cycle and pregnancy-rank, therefore, supporting our hypothesis. The results indicate that large variability exists in BCS, and BCS contributes substantially to the differences in live weight. The findings suggest that when predicting BCS from live weight consideration of these factors is required and different prediction equations needed. Adjustments for differences between BCS should consider age group, stage of the annual production cycle and pregnancy-rank. The relationships found between live weight and body condition score support the possibility of using live weight as a proxy for body condition score.

Chapter 7. Predicting ewe body condition score using lifetime live weight and liveweight change, and previous body condition score record

Chapter published as: **Semakula J, Corner- Thomas R.A, R, Morris S.T, Blair, H.T, Kenyon, P.R**. **(2020).** Predicting Ewe Body Condition Score Using Lifetime Live weight and Liveweight Change, and Previous Body Condition Score Record. *Animals*, 10(7):1182.

Abstract

Body condition score (BCS) in sheep (*Ovis aries*) is a widely used subjective measure of body condition. Body condition score and live weight have been reported to be statistically and often linearly related in ewes. Therefore, it was hypothesized that current BCS could be accurately and indirectly predicted using a ewe's lifetime live weight, liveweight change and previous BCS record. Ewes born between 2011 and 2012 (n = 11,798) were followed from 8 months to approximately 67 months of age in New Zealand. Individual ewe data were collected on live weight and body condition score at each stage of the annual production cycle (pre-breeding, pregnancy diagnosis, pre-lambing and weaning). Linear regression models were fitted to predict BCS at a given ewe age and stage of the annual production cycle using a ewe's lifetime live weight records (liveweight alone models). Further, linear models were then fitted using previous BCS and change in live weight in addition, to the lifetime live weight records (combined models). Using the combined models improved ($p < 0.01$) the R^2 value by 39.8% (from 0.32 to 0.45) and lowered the average prediction error by 10 to 12% (from 0.29 to 0.26 body condition scores). However, a significant portion of the variability in BCS remained unaccounted for (39 to 89%) even in the combined models. The procedures found in this study, therefore, may overestimate or underestimate measures by 0.23 to 0.32 BCS, which could substantially change the status of the ewe leading to incorrect management decisions. However, the findings do still suggest that there is potential for predicting ewe BCS from live weight using linear regression if key variables affecting the relationship between BCS and live weight are accounted for. This would benefit farmers by allowing for targeted nutritional management of individual animals to maximize overall flock productivity

7.1 Introduction

Body condition score (BCS) in sheep (Ovis aries) is a widely used subjective measure of the degree of body fatness (Jefferies, 1961; Russel et al., 1969; Morris et al., 2002; Vieira et al., 2015). It examines the degree of soft tissue coverage (predominantly fat and muscle) in lumbar region (Jefferies, 1961; Kenyon et al., 2014). Body condition score utilizes a 1.0−5.0 scale using half units or quarter units, and is undertaken by the palpation of the lumbar vertebrae (spinous and transverse process) immediately caudal to the last rib and above the kidneys (Kenyon et al., 2014). Unlike live weight (LW), BCS is not affected by factors such as variations in gut-fill, fleece weight, pregnancy and frame size that confound live weight as a measure of animal size to predict body condition (Coates and Penning, 2000b; Kenyon et al., 2014). Body condition score can be easily learned and is cost-effective and requires no specialist equipment (Kenyon et al., 2014). Knowledge of sheep BCS ensures that available feed resources are efficiently utilized, subtle differences in body condition not visibly noticeable are determined, there is instant awareness by producers about major changes in body fatness and the monitoring of trends in nutrition and body weight.

Even though using BCS offers several advantages over live weight (LW) to better manage flocks, farmers do not regularly use this technique. For example, while 96% of Australian producers indicated they monitored the body condition, only 7% conducted hands-on BCS (Jones et al., 2011). In New Zealand, 4% of farmers (Corner-Thomas et al., 2016) used BCS as a management tool. Farmers either rely on a visual inspection, which is inaccurate, or prefer to use live weight measures only (Besier and Hopkins, 1989). The reasons for low BCS adoption among farmers include: (1) the subjective nature of BCS, depending on assessor judgement; (2) being labor-intensive and (3) needs assessor training, that should be recalibrated over time (Kenyon et al., 2014). Strategies used to increase the use of BCS among farmers and its reliability included farmer training workshops and regular recalibration (Kenyon et al., 2014). However, given the apparently low rate of farmer uptake especially in large extensively managed flock systems, these strategies have been unsuccessful, likely due to not directly addressing how to reduce the labour burden with hands-on BCS. Therefore, it could be argued that, reliable and accurate alternative automated methods to estimate body condition score would be advantageous and would improve farmer uptake and use of BCS. Ideally, any automatic system to be utilized on extensive and intensive sheep farms would be based on a management tool already utilized on farms, to reduce workload and it would be quick and not subjective in nature.

The relationship between BCS and LW is documented in sheep (Kenyon et al., 2014; McHugh et al., 2019) with BCS being positively and generally linearly associated with live weight

(Kenyon et al., 2014). In Chapter 6, the factors affecting the relationship between BCS and LW in ewes such as age, stage of the annual production cycle and breed of ewe were assessed. Studies suggest correlations between BCS and LW can be between 0.20 to 0.89 and are stronger in mature ewes (r = 0.73 to 0.89) (Morel et al., 2016; McHugh et al., 2019). If the relationship between BCS and LW is predictable, then in theory, measurements of the latter could be used as predictors of BCS. In European sheep breeds, mature live weight occurs between 25 to 50 months of age (Wiener, 1967; Cake et al., 2006). Therefore, it could be postulated that, at approximately three years of age, when mature live weight is reached, a stable base BCS-LW relationship would be established. If this was indeed the case then, as a sheep ages further, future live weights, based on body condition score-live weight prediction equations could be used to predict a BCS or change in BCS with a fair degree of accuracy and reduce the need for hands-on BCS measurement.

In large extensive flock systems farmers regularly weigh sheep and increasingly more are using electronic tags (Corner-Thomas et al., 2016). Both conventional and modern weighing systems combined with individual electronic identification can now allow lifetime data to be collected more easily and quickly on large sheep flocks. Using this technology, combined with an individual BCS at a given point in their lifetime, therefore, can allow a specific stage of life BCS live weight relationship to be developed. Thus, using a set of established equations it should be possible to have a predicted BCS instantly calculated at each live weighing for each sheep. However, these have yet not been developed. If these could be developed, they could be incorporated into the electronic weigh heads of modern weigh systems to give farmers predictions of BCS. To date, this has not been tested. The aim of this study was to investigate the possibility of using lifetime live weight, liveweight change and previous BCS to predict a ewe's current body condition score.

7.2 Materials and Methods

7.2.1 Farms and animals used and data collection

The current study utilized data collected between 2011 and 2016 from two commercial New Zealand sheep farms (A and B) as part of normal routine farm management. All ewes (Romney breed) were weighed (to the nearest 0.1 kg) using static digital weighing scales (Tru-Test group, model XR5000). Body condition score was undertaken by experienced assessors using a 1.0-5.0 scale (1.0 = thin, 5.0 = obese) with sheep assessed to the nearest 0.5 of a BCS (Jefferies, 1961; Kenyon et al., 2014) at four time periods within an annual production cycle namely, pre-breeding, pregnancy diagnosis, pre-lambing and weaning. Data were collected over six individual years as ewes aged 8−18 to ≥67 months). A full description of the data used in the present study and sample characteristics is given in Chapter 6. In Chapter 6, the nature of association of the relationship between LW and BCS and the factors affecting this relationship were studied. The present study explores the possibility of utilizing the established relationship in the study above to indirectly predict a ewe's current BCS using previous live weight, liveweight change and BCS record. Table 7.1 below gives a summary of the variables used in BCS prediction models.

Table 7.1 Explanation of live weight, liveweight change and body condition score variables by ewe age group and stage of the annual production cycle.

Age (months)	Stage of the annual production cycle	*Live weight	§BCS	£Change in live weight
8–18	Pre-breeding	WM1	BM1	
	Pregnancy diagnosis	WP1	BP1	WT11(WP1–WM1)
	Pre-lambing	WL1	BL1	WT12(WL1–WP1)
	Weaning	WW1	BW1	WT13(WW1–WL1)
19–30	Pre-breeding	WM2	BM2	T2-T1(VM2–WW1)
	Pregnancy diagnosis	WP2	BP2	WT21(WP2–WM2)
	Pre-lambing	WL2	BL2	WT22(WL2–WP2)
	Weaning	WW2	BW2	WT23(WW2–WL2)
31–42	Pre-breeding	WM3	BM3	T3-T2(VM3–WW2)
	Pregnancy diagnosis	WP3	BP3	WT31(WP3–WM3)
	Pre-lambing	WL3	BL3	WT32(WL3–WP3)
	Weaning	WW3	BW3	WT33(WW3–WL3)
43–54	Pre-breeding	WM4	BM4	T4-T3(VM4–WW3)
	Pregnancy diagnosis	WP4	BP4	WT41(WP4–WM4)
	Pre-lambing	WL4	BL4	WT42(WL4–WP4)
	Weaning	WW4	BW4	WT43(WW4–WL4)
55-65	Pre-breeding	WM5	BM5	T5-T4(VM5–WW4)
	Pregnancy diagnosis	WP5	BP5	WT51(WP5–WM5)
	Pre-lambing	WL5	BL5	WT52(WL5–WP5)
	Weaning	WW5	BW5	WT53(WW5–WL5)
≥67	Pre-breeding	WM6	BM6	T6-T5(VM6–WW4)
	Pregnancy diagnosis	WP6	BP6	WT61(WP6–WM6)
	Pre-lambing	WL6	BL6	WT62(WL6–WP6)
	Weaning	WW6	BW6	WT63(WW6–WL6)

*Live weight; at pre-breeding (WM), pregnancy diagnosis (WP), pre-lambing (WL) and weaning (WW).
§BCS; at pre-breeding (BM), pregnancy diagnosis (BP), pre-lambing (BL) and weaning (BW).
£Change in live weight: WT; change in live weight between successive measurements within age groups, DT-T; change in live weight between successive measurements between age groups

7.2.2 Statistical analyses

Data were analyzed using R program version 3.3.4 (R Core Team, 2016) with package extensions in the caret package (Kuhn, 2008). It was not possible to observe a strict measurement collection protocol, therefore, missing values occurred in our dataset. To fill in the missing values, we used the preProcess function from the caret package in R (bagimput method). This method constructs a "bagging" model for each of the available variables based on regression trees, using all other variables as predictors while preserving the original data distribution structure(Kuhn, 2008). Live weight data were also normalized and centered during analysis using the same preProcess function above.

Body condition score data is both discrete and ordered in nature, which makes multi class classification regression approaches such as ordinal logistic or nominal regression more suitable for its analysis. However, when the underlying assumptions are grossly violated or when classes are extremely imbalanced (Leevy et al., 2018), classification statistical methods become less accurate (Tharwat, 2020). Triguero et al. (2015) categorizes class imbalances above 50:1 for any two outcomes as high-class imbalance. Strategies to overcome the challenge of class imbalance include oversampling, under-sampling and synthetic minority over sampling (Chawla et al., 2002). Such methods of circumventing class imbalances hold in cases of "reasonable" imbalance (Triguero et al., 2015). In case of high-class imbalance, the samples generated become less representative of the true sample distribution leading to under or over fitting the model. In the present study it was not possible to conduct classification regression using a full BCS scale (1.0−5.0) due to high class imbalance (1:1 to 1:280). The mitigation approaches to high-class imbalance may include modification of scale to a size that improves the distribution of values (not favourable for full scale prediction) or the use of other statistical methods robust to class imbalance such as multivariate (multiple regression) methods for interval and continuous data (Norman, 2010). In cattle, multivariate regression has successfully been used to predict BCS from physical body measurements and 3-D camera image data (Martins et al., 2020).Therefore, based on the previously outlined points, multivariate linear regression was used to predict ewe BCS from live weight.

7.2.3 Variable selection, model building and validation

Initially, the best predictor combinations for each BCS were selected through the regularization and variable selection technique implemented in R program (R Core Team, 2016) using the elastic net method in the glmnet extension (Friedman et al., 2010) in caret package (Kuhn, 2008). The elastic net method combines the power of two penalized-regularization methods (ridge and lasso regression) to search for the number of variables as well as handling collinearity (Archer and Williams, 2012).

All models were constructed, fitted and validated using algorithms, implemented in four steps. The steps included i) data partitioning, ii) resampling, iii) model training and iv) validation. Data partitioning involved dividing the initial dataset (with stratification preserving the class proportions) into training and testing datasets in a ratio of 3:1, with replacement. Resampling involved using bootstrapping and aggregation (Tropsha et al., 2003) procedures implemented in R (R Core Team, 2016) using caret package (Kuhn, 2008) to select 10 subsamples from the training set, and repeating the resampling three times. Model training involved fitting of the model using the training dataset subsamples (10) from which, nine were used for computing the parameters (i.e. β) while the remaining one part was used for error estimation (ε). Finally, all parameter estimates or probabilities from each sub sample were averaged to obtain the final value (estimate) with a 95% confidence interval.

Two multiple regression approaches were evaluated for the possibility of predicting BCS on a full scale namely, general linear model (LM) using the generalized least squares (GLS) and linear mixed effects model (LMM) procedure in nlme package (Pinheiro et al., 2018). The LM was selected for subsequent analysis the variance between LMM and GLS showed no significant difference ($p < 0.05$). Using selected best predictors for each BCS, LM regression equations were fitted to predict the current BCS using lifetime (present and previous) live weight records (liveweight alone models). Later, the models were modified by using previous liveweight change and BCS scores in addition to lifetime (combined models). Consequently forty-eight (48) regression equations were generated for BCS prediction, half of which were from using liveweight alone models and the remaining half from the combined models. Lifetime measurements refer to those ewe measurements taken at same and previous time points, whereas previous measurements only refer to those preceding the current one. Liveweight change was defined as sequential retrospective change in liveweight between individual time points.

7.2.4 Model performance evaluation

The calibration model performance (based on training dataset) was assessed using two metrics (Theil, 1958; Botchkarev, 2019) adjusted coefficient of determination (Adj. R^2) and the root mean square error (RMSE). The validation for each BCS prediction model was conducted using the testing dataset, each repeated 1000 times. Several metrics were considered when assessing the quality of BCS prediction models, including the coefficient of determination (r^2), bias, root mean squared error (RMSE), residual prediction deviation (RPD), and the ratio of performance to interquartile distance (RPIQ) (McDowell et al., 2012). The formulae used to compute the error metrics and coefficient of determination are in Chapters 4 and 6. The success of the predictions for individual samples was determined using the percent error (MAPE or RPE). The best model would have the highest Adj. R^2, r^2, RPD, and RPIQ, and the lowest RMSE and PE. In addition, RPD has been classified (Bellon-Maurel et al., 2010) into three different categories, weak prediction (RPD < 1.4), reasonable (1.4 < RPD < 2.0) and excellent (RPD > 2.0). In a similar manner (Nawar and Mouazen, 2017), RPIQ has been divided into four categories, very poor prediction (RPIQ < 1.4), fair (1.4 < RPIQ < 1.7), good (1.7 < RPIQ < 2.0), very good (2.0 < RPIQ < 2.5) and excellent (RPIQ > 2.5).

7.3 Results

7.3.1 Correlation between all BCS and live weights

There was association between live weight and BCS in all age groups and stage of the annual production cycle, but the association was characterized as being weak to moderate (Tables 7.2 and 7.3). The relationships, however, were stronger when live weight and BCS measurements were from the same time point ($0.25 \leq r \leq 0.67$), compared with when lifetime (i.e. including the same time point and previous) records were used ($-0.18 \leq r \leq 0.67$). In terms of stage of the annual production cycle, the correlation was strongest at weaning ($-0.08 \leq r \leq 0.67$) and weakest pre-lambing ($-0.18 \leq r \leq 0.49$).

7.3.2 Linear regression (coefficient of determination (R^2) and number of predictors)

To predict current BCS, all current and previous individual live weights (liveweight alone models) were included in linear regression equations (Appendix VIII Tables 1a and 1b for liveweight alone models and Appendix VIII Tables 2a and 2b for combined models). Across age groups, the change in adjusted R^2 value showed no clear pattern (Figure 7.1). The adjusted R^2 values averaged 0.32 and did not get above 0.49, regardless of time point. There was no trend for R^2 to improve at older ages, when a greater amount of previous live weight information was known. It was observed that in general the adjusted R^2 value was highest at weaning but lowest at pre-lambing.

The average number of live weight predictors (significant variables) for BCS prediction was seven (1 to 16) with no clear pattern of change over time. To improve the prediction of current BCS, a combination of all preceding BCS, and prior live weights and their sequential retrospective differences (change in live weight between individual time points) were included in the regression equations (combined models) and are shown in (Appendix VIII Tables 2a and 2b). The number of significant predictors for BCS was higher (average: 25, from 1 to 59) in the combined models compared with liveweight alone models.

The adjusted R^2 values (Figure 7.1) for ewe BCS prediction ranged from 0.11 to 0.61 for liveweight alone models or combined models. Although, there was no clear trend for R^2 improvement with age, it appeared to be affected by stage of the annual production cycle. Notable was the generally low R^2 value at pre-lambing in both combined and liveweight alone models. The adjusted R^2 increased with the number of variables in the combined model in a similar manner to the liveweight alone models. Using more predictors in addition to live weight, increased the adjusted R^2 value by 39.8. % (from 32.5 to 45.4%) or 1.4 times and the number of significant predictors at each stage of the annual production cycle by 3.6 (average number of variables for combined models liveweight alone models divided by average number of variables for liveweight alone models) times. A significant portion of the variability in BCS remained unaccounted for (38 − 89%) in the combined models, with some of the initial live weight variables in the liveweight alone models being considered non-significant ($p > 0.05$) in the combined models.

Table 7.2 Correlation coefficients between individual live weight and body condition scores across stage of the annual production cycle in ewes between 8 and 42 months

		Body condition score											
Weight	n	BM1	BP1	BL1	BW1	BM2	BP2	BL2	BW2	BM3	BP3	BL3	BW3
WM1	11,798	0.38	0.13	0.13	−0.05ns	0.00ns	0.08	−0.12	0.18	0.02ns	0.09	0.01ns	0.19
WP1	11,124	0.32	0.36	0.46	0.11	0.00ns	0.10	−0.02ns	0.16	0.05	0.08	0.03ns	0.22
WL1	8,074	0.28	0.18	0.49	0.25	−0.11	0.16	0.43	0.21	0.21	0.18	0.08	−0.04
WW1	8,499	0.09	0.25	0.44	0.67	0.41	0.28	0.33	0.12	0.17	0.11	0.06	0.04
WM2	8,393	0.12	0.25	0.33	0.54	0.49	0.25	0.26	0.15	0.16	0.14	0.09	0.11
WP2	7,991	0.14	0.36	0.29	0.25	0.37	0.39	0.01ns	0.15	0.04	0.11	0.14	0.30
WL2	5,362	0.15	0.34	0.45	0.41	0.29	0.40	0.25	0.11	0.10	0.14	0.07	0.15
WW2	6,950	0.13	0.28	0.33	0.25	0.19	0.21	0.11	0.53	0.39	0.32	0.26	0.29
WM3	6,651	0.14	0.06	0.12	0.20	0.16	0.24	0.21	0.48	0.51	0.45	0.29	0.21
WP3	6,308	0.16	0.13	0.29	0.26	0.13	0.31	0.29	0.46	0.43	0.51	0.32	0.19
WL3	2,700	0.13	0.17	0.15	0.20	0.13	0.25	0.24	0.33	0.38	0.45	0.32	0.16
WW3	5,579	0.12	−0.03ns	0.01ns	0.09	0.12	0.21	0.10	0.38	0.23	0.32	0.26	0.60
WM4	5,149	0.12	−0.04	0.02ns	0.11	0.12	0.22	0.16	0.32	0.24	0.32	0.24	0.43
WP4	4,944	0.14	−0.11	0.01ns	0.13	0.08	0.27	0.30	0.34	0.27	0.39	0.27	0.34
WL4	3,224	0.12	−0.03ns	0.02ns	−0.03ns	0.09	0.22	0.13	0.34	0.18	0.31	0.19	0.37
WW4	4,440	0.06	0.06	0.06	0.17	0.11	0.13	0.09	0.19	0.17	0.18	0.15	0.21
WM5	4,314	0.07	−0.03ns	−0.02ns	0.11	0.06	0.14	0.15	0.21	0.19	0.22	0.18	0.15
WP5	4,146	0.09	−0.07	0.01ns	0.16	0.05	0.16	0.25	0.20	0.20	0.25	0.18	0.11
WL5	2,677	0.10	−0.11	0.02ns	0.19	0.02ns	0.15	0.21	0.16	0.20	0.20	0.08	0.03
WW5	2,695	0.08	−0.15	0.01ns	0.15	0.03ns	0.16	0.27	0.23	0.22	0.22	0.08	0.08
WM6	1,437	0.09	−0.15	−0.06	0.12	−0.02ns	0.13	0.23	0.16	0.16	0.22	0.10	0.06
WP6	1,334	0.09	−0.12	−0.05	0.13	−0.04	0.15	0.28	0.15	0.16	0.23	0.10	0.01ns
WL6	879	0.08	0.09	0.02ns	0.11	0.01ns	0.02ns	0.05	0.08	0.08	0.15	0.09	0.11
WW6	563	0.06	−0.03ns	−0.03ns	0.11	−0.03ns	0.01ns	0.05	0.10	0.09	0.08	0.09	0.09

BM, BP, BL, BW indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing and at weaning respectively. WM, WP, WL, WW indicate live weight prior to pre-breeding, at pregnancy diagnosis, prior to lambing and at weaning respectively. Grey shade (major diagonal) indicates live weights and BCS correlation coefficient values from the same time point. n: indicates sample size. ns: superscript indicates no significance at $p < 0.05$.

Table 7.3 Correlation coefficients between individual live weight and body condition scores across stage of the annual production cycle in ewes above 42 months of age

		Body condition score											
Weight	n	BM4	BP4	BL4	BW4	BM5	BP5	BL5	BW5	BM6	BP6	BL6	BW6
WM1	11,798	0.03ns	−0.05	0.3	0.11	0.18	0.06ns	−0.03ns	−0.03ns	−0.09	−0.09	−0.03ns	0.01ns
WP1	11,124	0.02ns	−0.05	0.33	0.13	0.19	0.05ns	−0.04ns	−0.05	−0.11	−0.10	0.00ns	0.04
WL1	8,074	0.20	0.10	−0.11	−0.04	−0.03ns	0.15	0.16	0.21	0.35	0.36	0.11	0.14
WW1	8,499	0.13	0.09	−0.18	0.04	0.01ns	0.04ns	0.10	0.12	0.15	0.11	0.03ns	0.06
WM2	8,393	0.14	0.08	0.01	0.07	0.10	0.07	0.07	0.08	0.09	0.08	0.06	0.08
WP2	7,991	0.04	0.00ns	0.43	0.21	0.29	0.09	−0.04ns	−0.08	−0.17	−0.15	0.12	0.04
WL2	5,362	0.11	0.10	0.01ns	0.11	0.17	0.09	0.04ns	0.05	0.06	0.05	0.10	0.10
WW2	6,950	0.13	0.06	0.30	0.19	0.25	0.10	0.04ns	0.03ns	−0.07	−0.06	0.06	0.09
WM3	6,651	0.23	0.11	0.20	0.12	0.20	0.14	0.11	0.13	0.10	0.11	0.11	0.15
WP3	6,308	0.25	0.17	0.26	0.13	0.22	0.21	0.15	0.15	0.14	0.18	0.12	0.18
WL3	2,700	0.22	0.14	0.09	0.12	0.20	0.14	0.09	0.15	0.10	0.14	0.12	0.15
WW3	5,579	0.47	0.29	0.38	0.19	0.27	0.18	0.11	0.14	0.12	0.15	0.06	0.20
WM4	5,149	0.53	0.35	0.33	0.17	0.27	0.22	0.16	0.16	0.17	0.16	0.10	0.22
WP4	4,944	0.51	0.46	0.33	0.12	0.24	0.29	0.21	0.23	0.27	0.30	0.14	0.21
WL4	3,224	0.32	0.17	0.46	0.20	0.32	0.23	0.10	0.11	0.07	0.11	0.07	0.13
WW4	4,440	0.26	0.18	0.18	0.55	0.40	0.31	0.23	0.20	0.15	0.16	0.09	0.22
WM5	4,314	0.26	0.16	0.19	0.30	0.48	0.39	0.28	0.20	0.25	0.26	0.15	0.29
WP5	4,146	0.28	0.22	0.13	0.20	0.32	0.48	0.33	0.26	0.31	0.33	0.17	0.22
WL5	2,677	0.24	0.16	0.03ns	0.12	0.16	0.20	0.31	0.26	0.24	0.25	0.03ns	0.16
WW5	2,695	0.27	0.15	0.05	0.12	0.14	0.24	0.29	0.63	0.45	0.39	0.20	0.25
WM6	1,437	0.28	0.15	0.03ns	0.06	0.14	0.18	0.25	0.38	0.59	0.49	0.24	0.32
WP6	1,334	0.24	0.15	0.04	0.03ns	0.07	0.21	0.19	0.33	0.48	0.56	0.25	0.28
WL6	879	0.17	0.07	0.05	0.15	0.22	0.13	0.09	0.25	0.25	0.34	0.28	0.28
WW6	563	0.16	0.04	0.07	0.13	0.16	0.11	0.01ns	0.19	0.24	0.26	0.27	0.64

BM, BP, BL, BW indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing and at weaning respectively. WM, WP, WL, WW indicate live weight prior to pre-breeding, at pregnancy diagnosis, prior to lambing and at weaning respectively. Grey shade (major diagonal) indicates live weights and BCS correlation coefficient values from the same time point. n: indicates sample size. ns: superscript indicates no significance at $p < 0.05$.

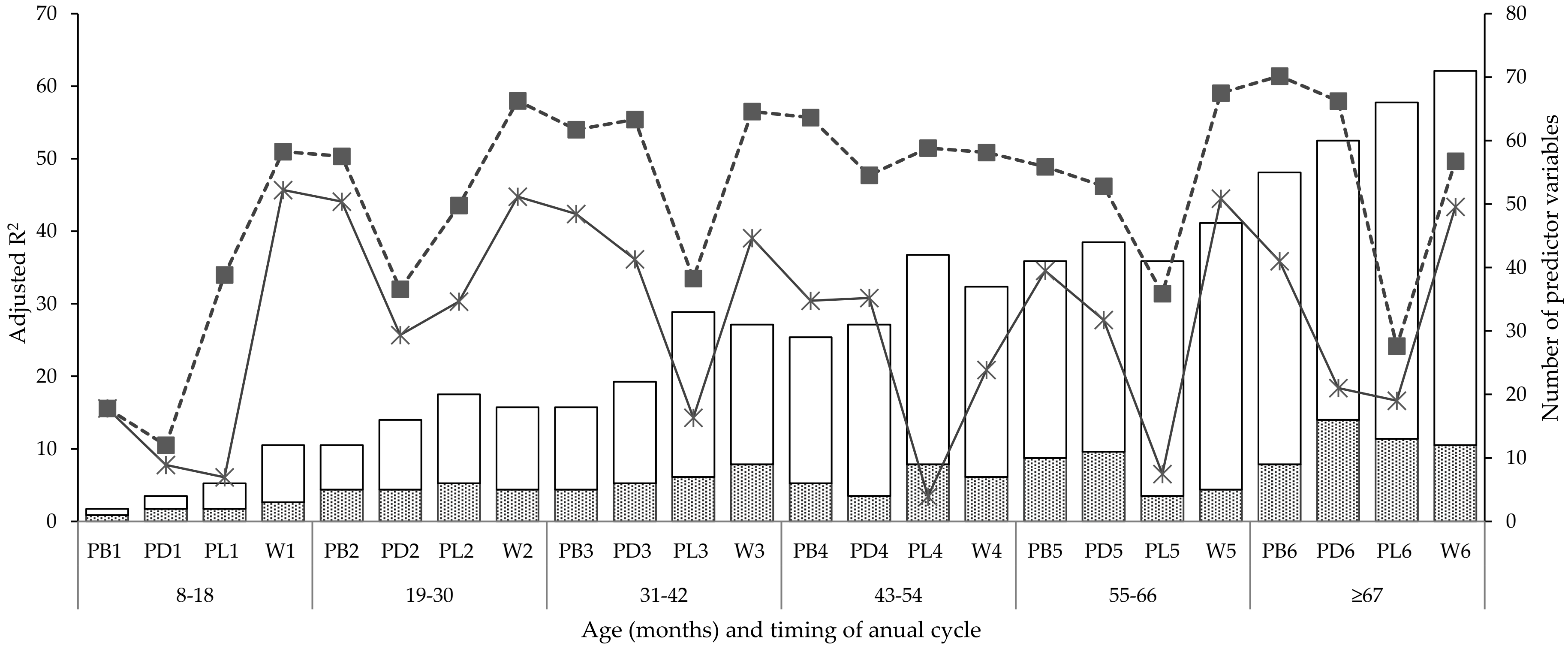


Figure 7.1 Adjusted R^2 (solid line: liveweight alone models, dashed: combined models) and number of predictors (dotted bar: liveweight alone models and white bar: combined models) for BCS prediction across the stage of the annual production cycle and ewe age group. BM, BP, BL, BW indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing and at weaning respectively.

7.3.3 *Prediction error metrics*

The BCS model prediction error metrics (MAE, RMSE, MAPE, RPE) and r^2 varied within across age group and stage of the annual production cycle when live weight or combined models were used to predict BCS (Tables 7.5 and 7.6). The average prediction error associated with BCS prediction in live weight and in the combined models in terms of MAE and RMSE was 0.26 (0.23 to 0.32) and 0.32 (0.28 to 0.41) body condition scores, respectively. The magnitude of the error values was categorized as being moderate to high in both live weight and in the combined models, given the scale of measurement and smallest unit of measurement (0.5). The BCS predictions using the liveweight alone models were, on average, 9.3 (7.60 to 11.50) to 11.6 (9.50 to 14.62) % from the actual value. The models were categorized as weak (RPD: 1.02 to 1.39) or very poor to fair (RPIQ: 1.28 to 1.79).

The model prediction error metrics for the combined models varied across age group and stage of the annual production cycle but were significantly ($p < 0.01$) reduced compared with the liveweight alone models. The average prediction error associated with BCS prediction using the combined models in terms of MAE and RMSE was reduced by 0.04 (10% to 12%) body condition scores. Overall, the combined models improved BCS prediction from weak to reasonable (PRD: 1.40) or good (RPID: 1.75). The results also showed positive and negative bias for both models, and indication of tendency to underestimate or overestimate BCs measurement.

Table 7.4 Coefficient of determination (r^2), bias, root mean square error (RMSE), residual prediction deviation (RPD) and the ratio of performance to interquartile distance (RPIQ) based on testing data, for the prediction of BCS in ewes between 8 and 42 months by stage of the annual production cycle using live weight (liveweight alone models) and a combination of predictors (combined models).

	Age group											
	8–18				19–30				31–42			
	BM1	BP1	BL1	BW1	BM2	BP2	BL2	BW2	BM3	BP3	BL3	BW3
BCS range	1.5–4.5	1.5–4.5	1.5–4.0	1.5–4.5	1.5–5.0	1.5–4.0	1.5–4.0	1.5–5.0	1.5–4.5	1.5–4.0	1.5–4.0	1.0–4.5
	Liveweight alone models[a]											
r^2 %	15.7	9.1	6.1	45.4	39.4	22.6	26.9	43.7	42.2	24.1	12.4	40.1
Bias	0.01	0.002	−0.01	0.00	0.01	−0.01	0.00	0.00	0.00	0.00	0.01	0.01
MAE	0.30	0.31	0.32	0.27	0.24	0.24	0.25	0.30	0.23	0.24	0.28	0.26
RMSE	0.38	0.43	0.38	0.53	0.27	0.30	0.32	0.38	0.28	0.31	0.35	0.33
MAPE	11.11	13.15	10.54	9.27	11.06	9.11	9.33	10.78	8.29	8.39	9.77	8.94
RPE	12.89	14.4	12.36	12.12	11.76	11.39	11.95	13.66	10.09	10.84	12.12	11.35
RPD	1.12	1.06	1.03	1.36	1.20	1.14	1.16	1.22	1.28	1.23	1.09	1.31
RPIQ	1.32	1.28	1.47	1.47	1.52	1.67	1.56	1.32	1.79	1.61	1.43	1.52
	Combined models[b]											
r^2 %	15.7	10.8	35.2	50.0	50.3	34.0	41.2	58.9	53.6	55.5	32.3	56.7
Bias	0.01	0.00	−0.01	−0.01	0.004	0.00	−0.01	−0.01	−0.003	0.00	0.001	−0.01
MAE	0.30	0.02	0.23	0.25	0.22	0.22	0.21	0.24	0.19	0.20	0.31	0.23
RMSE	0.38	0.02	0.28	0.32	0.28	0.28	0.28	0.31	0.24	0.26	0.24	0.29
MAPE	11.11	2.47	8.35	8.92	7.85	8.36	7.84	8.66	6.849	7.21	8.4	7.926
RPE	12.89	2.47	10.17	11.41	9.98	10.64	10.45	11.19	8.65	9.37	10.85	9.99
RPD	1.12	1.19	1.23	1.43	1.43	1.23	1.31	1.55	1.47	1.36	1.22	1.51
RPIQ	1.32	1.50	1.78	1.56	1.79	1.79	1.78	1.62	2.08	1.92	1.61	1.72

BM, BP, BL, BW indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing and at weaning respectively. Interpretation of measures: The best model has the highest r^2, RPD, and RPIQ, and the lowest RMSE and RPE. Ranges for values: r^2 (0: indicates that the model accounts for none of the variability of the response data around its mean, 1.0 indicates that the model accounts for all the variability). RPD (< 1.4: weak, 1.4 < RPD < 2.0: reasonable, > 2.0: excellent). RPIQ (< 1.4: very poor, 1.4 < RPIQ < 1.7: fair, 1.7 < RPIQ < 2.0: good, 2.0 < RPIQ < 2.5: very good, > 2.5: excellent). [a] Liveweight alone models indicates all previous and current weight. [b] Combined models indicates all previous and current weights, liveweight change and previous BCS. Bias (Positive value indicates overestimation; negative sign indicates underestimation).

Table 7.5 Coefficient of determination (r^2), bias, root mean square error (RMSE), residual prediction deviation (RPD) and the ratio of performance to interquartile distance (RPIQ) based on testing data, for the prediction of BCS in ewes above 42 months of age by stage of the annual using live weight (liveweight alone models) and a combination of predictors (combined models).

	Age group											
	43–54				55–66				≥67			
	BM4	BP4	BL4	BW4	BM5	BP5	BL5	BW5	BM6	BP6	BL6	BW6
BC range	1.0–4.0	1.0–4.0	1.5–4	1.5–4.0	1.0–4.0	1.0–4.0	2.0–4.0	1.0–4.0	1.5–4.0	1.5–4.5	1.5–3.5	1.5–4.5
	Liveweight alone models[a]											
r^2	37.5	32.1	15.3	40.2	33.7	25.9	15.1	42.4	34.9	36.2	12.6	41.8
Bias	−0.004	0.01	0.01	0.01	−0.01	−0.01	0	−0.02	0.01	−0.02	−0.01	−0.01
MAE	0.25	0.24	0.24	0.24	0.24	0.24	0.26	0.27	0.24	0.31	0.25	0.27
RMSE	0.31	0.31	0.32	0.31	0.29	0.33	0.32	0.34	0.31	0.38	0.32	0.34
MAPE	8.28	8.30	8.90	9.05	10.03	8.29	9.21	10.38	7.86	9.80	9.61	9.69
RPE	10.26	10.71	11.87	11.68	12.67	11.4	11.33	13.03	10.15	14.66	11.75	12.2
RPD	1.27	1.21	1.26	1.30	1.13	1.14	1.02	1.32	1.34	1.13	1.06	1.39
RPIQ	1.61	1.56	1.55	1.61	1.39	1.51	1.56	1.40	1.61	1.32	1.56	1.47
	Combined models[b]											
r^2	52.6	51.3	52.3	47.9	52.4	49.5	27.8	58.3	63.2	65.4	33.9	43.0
Bias	−0.003	−0.007	−0.013	0.012	0.002	0.009	−0.014	−0.001	0.011	−0.001	0.004	−0.007
MAE	0.21	0.20	0.22	0.22	0.22	0.20	0.22	0.22	0.20	0.22	0.23	0.30
RMSE	0.26	0.26	0.29	0.28	0.29	0.25	0.28	0.28	0.25	0.27	0.28	0.3756
MAPE	6.94	6.9	8.19	8.28	8.30	6.89	7.78	8.35	6.53	6.75	8.52	10.68
RPE	8.59	8.97	10.42	10.55	10.56	8.62	9.89	10.62	8.17	8.28	10.84	13.17
RPD	1.48	1.42	1.47	1.38	1.53	1.42	1.16	1.53	1.61	1.71	1.25	1.31
RPIQ	1.92	1.92	1.79	1.79	1.79	2.00	1.79	1.79	2.00	1.85	1.79	1.35

BM, BP, BL, BW indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing and at weaning respectively. Interpretation of measures: The best model has the highest r^2, RPD, and RPIQ, and the lowest RMSE and RPE. Ranges for values: r^2 (0: indicates that the model accounts for none of the variability of the response data around its mean, 1.0 indicates that the model accounts for all the variability). RPD (< 1.4: weak, 1.4 < RPD < 2.0: reasonable, > 2.0: excellent). RPIQ (< 1.4: very poor, 1.4 < RPIQ < 1.7: fair, 1.7 < RPIQ < 2.0: good, 2.0 < RPIQ < 2.5: very good, > 2.5: excellent). [a] Liveweight alone models indicates all previous and current weight. [b] Combined models indicates all previous and current weights, liveweight change and previous BCS. (Positive value indicates overestimation; negative sign indicates underestimation).

7.4 Discussion

The aim of this study was to explore the possibility of predicting BCS from lifetime live weight, liveweight change and previous BCS over time in ewes as they aged from eight through to approximately sixty-seven months. This appears to be the first study to attempt this in sheep. Previous studies have examined the relationship between live weight and BCS at a given time point (Sezenler et al., 2011; McHugh et al., 2019).

This study demonstrated that BCS prediction models based on a ewe's live weight record or a combination of live weight, liveweight change and previous BCS improved the proportion of variability in current BCS accounted for above that observed in Chapter 6. Further, it was demonstrated that despite BCS and LW being linearly correlated (Kenyon et al., 2014; McHugh et al., 2019), the relationship is weak when predicting using linear regression, even in older individuals which would have attained maturity. The results also indicated that the role of prior live weight measurements in predicting BCS diminishes as the time gap between measurements points increased. This indicates that using early life live weights alone would likely be unreliable in predicting future BCS. Further, the effect of liveweight change on BCS prediction was more significant during the early years of a ewe than in her later years, which implies that, liveweight change may cease to be an important predictor of BCS after maturity is reached.

The variability in BCS explained for both live weight and combined models increased with the number of predictors in the model. This was expected as it is known that as the number of predictors that significantly relate with the dependent variable increase, the proportion of the variance due to the regression increases (Li, 2017). However, in this study, a considerable amount of variability in BCS ($0.58 \leq R^2 \leq 0.91$ and $0.39 \leq R^2 \leq 0.89$) remained unaccounted for in both liveweight alone and combined models, respectively. Potential reasons for the apparent failure for both liveweight alone and combined models to account for more of the variability in BCS include BCS binning (due to not being a continuous variable), assessor consistency over time, losses in live weight due to gut-fill and urination when ewes are weighed at different times, fleece weight and wetness, and confounding of live weight with conceptus weight. The consistency of the BCS data can vary between (5 to 27% and 40 to 60%) and within (16 to 44% and 60 to 90%) operator for inexperienced and experienced assessors, respectively (Yates and Gleeson, 1975; Kenyon et al., 2014). Liveweight losses resulting from fluctuations in the gut-fill can account for between 5 and 23% of total live weight in ruminants (Hughes, 1976; Semakula et al., 2019). Thus, when an individual's live weight is recorded with respect to when the animal was fed, can influence the accuracy of a live weight. The present study did not measure for individual time off feed prior to weighing, a function that many electronic weighing systems have the potential to do. As the pregnancy advances, conceptus weight increases depending on the number of fetuses carried (Kleemann and Walker, 2005), which could have affected the live weight and BCS differently. The present dataset did not have information on the individual stage of pregnancy for each ewe. Future studies should examine if the accuracy of the prediction can be ameliorated by incorporating these two variables. In regression, the independent variable measurement is assumed to be measured with high precision, thus, it is not expected to contribute to residual error (Dosne et al., 2016). Therefore, losses in live weight due to gut-fill changes and urination in relation to when ewes are weighed at different times and the effect of pregnancy on live weight are of concern, as they affect live weight which is an independent variable for BCS prediction. When independent variables are not exact, estimations based on the standard assumption leads to inconsistent parameter estimates even in very large samples (Hausman, 2001; Pischke, 2007). Thus, if errors in the measurement of live weight could be minimized, then the resulting error term in the regression could all be attributed to BCS measurement, which should improve the model goodness-of-fit and accuracy. In order to reduce this measurement error, it would be imperative that liveweight losses due to delayed weighing be accounted for with respect to time of delay (period from when the animal last fed to weight recording) in using prediction equations. Time-dependent, live weight adjusting equations for ewes have been developed but are not regularly used (Wishart et al., 2017).

In the present study, the prediction models using liveweight alone had large error (MAE and RMSE) and low RPD and RPIQ values, which led to high error rates. Combined models reduced the magnitude of all the prediction error metrics to near acceptable levels. Although error (MAPE) up to 20% is acceptable for setting dosage rates in the veterinary pharmaceutical industry (Leach and Roberts, 1981), error of more than 10% can be problematic (Alexander et al., 2015; Hagerman et al., 2017; Lalic et al., 2018) in other agricultural filed MAPE or RMSPE. In this study, values were approximately 9 to 12% for liveweight alone models and 8 to 10% for the combined models. The moderate to large error values (one-half to two-thirds of the smallest unit on a 0.5 decimal scale) in BCS prediction in the present study (where a 0.5-unit change in BCS changes the performance-rank of a ewe) could greatly influence management decisions. In theory, both models should have had resolutions of approximately 0.02 (maximum span = 0.5 / smallest possible increment = (2) ^ maximum range of possible values) body condition score. However, due to the rigid nature (discrete or noncontinuous scale with no values in between the fixed points) of the scale used, such resolutions are not achievable. It has been suggested that decisions concerning targeted feeding and management of ewes to maximum performance are based on a minima BCS (i.e. 2.5) or a critical range of BCS values (i.e. 2.5 to 3.5) (Kenyon et al., 2014). The predictions found in this study may, therefore, overestimate or underestimate measures by 0.23 to 0.32 BCS, which could substantially change the status of the ewe leading to incorrect management decisions, which in turn could reduce flock productivity.

7.5 Conclusion

Applying a ewe's live weight record or a combination of live weight, liveweight change and previous BCS increased the proportion of variability in current BCS accounted for above that observed in Chapter 6. This improvement in the proportion of variability in BCS accounted for was greatest in the combined models. Further, the BCS prediction accuracy metrics across age groups and stage of the annual production cycle and over time (years) were greater in combined models compared with the liveweight alone models. This indicates that BCS could be better predicted if additional variables (live weight, liveweight change and previous BCS) were included in the multiple regression equation rather than lifetime liveweight alone. These relationships could potentially be incorporated in electronic weighing systems that utilize lifetime data. This would be especially useful when applied to large extensively run sheep flocks. However, a significant portion of the variability in BCS remained unaccounted for (39 to 89%) even in the combined models. It is possible that the prediction models could be improved if additional information such as stage of pregnancy, number of foetuses carried, and time off feed were utilized and warrants further investigation.

Chapter 8. Application of machine learning algorithms to predict body condition score from live weight records of mature Romney ewes

Chapter published as: **Semakula J, Corner- Thomas R.A, R, Morris S.T, Blair, H.T, Kenyon, P.R**. **(2021).** Application of Machine Learning Algorithms to Predict Body Condition Score from Liveweight Records of Mature Romney Ewes. *Agriculture*. 2021; 11(2):162.

Abstract

Body condition score (BCS) in sheep (*Ovis aries*) is a widely used subjective measure of the degree of soft tissue coverage. Body condition score and live weight are statistically related in ewes; therefore, it was hypothesized that BCS could be accurately predicted from live weight using machine learning models. Individual ewe live weight and body condition score data at each stage of the annual production cycle (pre-breeding, pregnancy diagnosis, pre-lambing and weaning) at 43 to 54 months of age were used. Nine machine learning (ML) algorithms including Ordinal logistic regression, Multinomial regression, Linear Discriminant Analysis, Classification and Regression Tree, Random Forest, K-Nearest Neighbours, Support Vector Machine, Neural Networks and Gradient boosting decision trees) were applied to predict BCS from a ewe's current and previous live weight record. A three class BCS (1.0−2.0, 2.5−3.5, >3.5) scale, was used due to high-class imbalance in the 5-scale BCS data. The results showed that using ML to predict ewe BCS at 43 to 54 months of age from current and previous live weight could be achieved with high accuracy (> 85%) across all stages of the annual production cycle. The gradient boosting decision tree algorithm (XGB) was the most efficient for BCS prediction regardless of season. All models had balanced specificity and sensitivity. The findings suggest that there is potential for predicting ewe BCS from live weight using classification machine learning algorithms.

8.1 Introduction

Body condition score (BCS) in sheep (*Ovis aries*) is a widely used subjective measure of the degree of soft tissue coverage (predominantly fat and muscle) of the lumbar vertebrae region (Jefferies, 1961; Kenyon et al., 2014). Body condition score is based on a 1.0−5.0 scale using half units or quarter units and is conducted by palpation of the lumbar vertebrae immediately caudal to the last rib above the kidneys (Kenyon et al., 2014). Unlike live weight (LW), BCS is not affected by fluctuations in gut-fill, fleece weight, and frame size which confound live weight as a measure of animal size to give an indication of body condition (Coates and Penning, 2000b). BCS can be easily learned and is cost-effective and requires no specialist equipment (Kenyon et al., 2014). The optimal BCS range for ewe performance is 2.5 to 3.5 (Kenyon et al., 2014), outside this range performance is either adversely affected or it is inefficient in terms of performance per kg of feed eaten (Morel et al., 2016). Farmers can use targeted feeding based on this optimal range to optimise overall performance.

Despite the advantages of using BCS over live weight (LW) for flock management, many farmers in extensive farming systems do not regularly do so. For instance, only 7% and 40% of the farmers indicated that they conducted hands-on BCS in Australia and New Zealand, respectively (Jones et al., 2011; Corner-Thomas et al., 2016). Farmers often rely on visual inspection, a method which is inaccurate, or only use live weight measure (Besier and Hopkins, 1989) which is influenced by factors including gut-fill variation, frame size, physiological stage and fleece weight (Kenyon et al., 2014). The low uptake of BCS among farmers may in some part be due to challenges such as assessor subjectivity and extra labour requirements (Kenyon et al., 2014). Attempts to increase the uptake of BCS among farmers, including use of promotional training workshops and regular training, have not yielded the desired outcome, likely because they do not directly alleviate the labour burden related to hands-on BCS (Kenyon et al., 2014). Therefore, accurate and reliable alternative methods to estimate body condition score with less hands-on measurement would be advantageous and would likely improve the uptake of BCS technology, especially for large flocks.

Ewe BCS and LW are correlated (Kenyon et al., 2014; McHugh et al., 2019). This relationship varies by age, stage of the annual production cycle (Chapter 6), and breed of animal (McHugh et al., 2019). In Chapter 6, it was reported that in Romney ewes, both LW and BCS plateaued after they reached 43−54 months of age, thereby establishing a stable base BCS-LW relationship. This means that, as a ewe ages, future live weights, based on BCS-LW prediction equations could potentially be used to predict a BCS with a degree of accuracy and reduce the need for hands-on BCS measurement.

Modern automated weighing systems with individual electronic identification, offer an opportunity to collect lifetime data relatively easily and quickly. With such large datasets, it has become possible to process and extract valuable information. In Chapter 6, multivariate regression models were applied to predict ewe BCS from lifetime live weight data as a ewe aged from eight to sixty-seven months. At best, these multivariate models explained 49% and 21% of the variability in BCS using the 5-scale (9 points) and 3-scale (3 points), respectively. Further, BCS was skewed with little variability due to the limited nature of the BCS scale used (1.0−5.0, in increments of 0.5). Using only discrete values such as BCS can lead to the heaping or grouping of all possible values (i.e. noncontinuous) at isolated points, affecting the resolution and ultimately the accuracy of any prediction model.

Approaches that circumvent the challenges of considering discrete as continuous data are required for BCS prediction. Classification-based models are recommended for discrete and categorical data analysis (Blaikie, 2003; Wicker, 2006; Sullivan and Artino, 2013; Bishop and Herron, 2015). Among these classification approaches, machine learning (ML) classification models have been used with greater success compared with traditional statistical methods in sheep production for early estimation of the growth and quality of wool in adult Australian merino sheep (Shahinfar and Kahn, 2018) and sheep carcass traits (Shahinfar et al., 2019) from early-life data. Machine learning utilizes algorithms whose logic can be learned directly from unique patterns in the data or inexplicitly through pre-programmed classical statistical methods (Khaledian and Miller, 2020). The successful use of ML algorithms in various fields of science, warrants their application in animal production problem-solving (Morota et al., 2018; Bakoev et al., 2020). Ideally, it should be possible to install this computer acquired intelligence into modern weighing systems to automatically explore patterns in lifetime live weights and predict BCS. The aim of this study was to investigate the use of machine learning algorithms to predict ewe BCS from current and previous live weight data. In the present study, ewe BCS was predicted for the ewes in their fourth year of life (43−54 months) at four stages of the annual system using previous live weight measurements.

8.2 Materials and Methods

8.2.1 Farms and animals used and data collection

The current study was a follow-up of the previous two studies (Chapter 6 and 7). In Chapter 6, only, the nature of the relationship between LW and BCS (linear) and the factors affecting their relationship (ewe age, stage of annual production cycle and pregnancy-rank) were determined. In Chapter 7, the potential of predicting ewe BCS as a continuous variable from live weight and previous BCS records was demonstrated. The resulting linear models had high prediction error (> 10%) and a greater part of the variability in BCS (from 39 to 89%) remained unexplained. The current study attempts to predict BCS from LW records in a more precise way, using machine learning algorithms. The details on how the animals were managed and data were collected were reported in Chapter 6.

8.2.2 Statistical analyses

Data were analyzed using R program version 4.3.4 (R Core Team, 2016) with caret package extensions (Kuhn, 2008). Data were initially explored to identify completeness and were summarized by BCS to determine class distribution. Missing values (n = 26) were imputed using the bagimput function from the caret package. This method constructs a "bagging" model for a given variable based on regression trees, using all other variables as predictors while maintaining the original data distribution structure (Kuhn, 2008). Live weight data were normalized and centered during analysis using the preProcess function from the caret package. The distribution of BCS at all stages of the annual production cycle showed that on a full BCS scale (1.0−5.0) there were high-class imbalances (more than 1:50 for any two classes) (Figure 8.1a and 8.1b). The average ratios of the class frequencies (minimum: maximum) were 1:216, 1:1336, 1:498 and 1:97 for pre-breeding, pregnancy diagnosis, pre-lambing and weaning, respectively (Figure 8.1a). The high-class or extreme imbalance was due to too few extreme BCS cases with the majority of individual BCS measurements ranging from 2.5 to 3.5.

Triguero et al. (2015) categorized class imbalances above 50:1 for any two outcomes as high-class imbalance. Body condition score data is both discrete and ordered in nature, which makes multi-class classification regression approaches more suitable for its analysis. However, when the underlying assumptions are grossly violated or when classes are extremely imbalanced (Leevy et al., 2018), classification statistical methods become less accurate (Tharwat, 2020). Strategies to overcome the challenge of class imbalance may include re-sampling techniques such as oversampling, under-sampling and synthetic minority oversampling (Chawla et al., 2002). Such methods of circumventing class imbalances hold in cases below 50:1 imbalance. In case of high-class imbalance, the samples generated become less representative of the true sample distribution leading to under- or over-fitting the model.

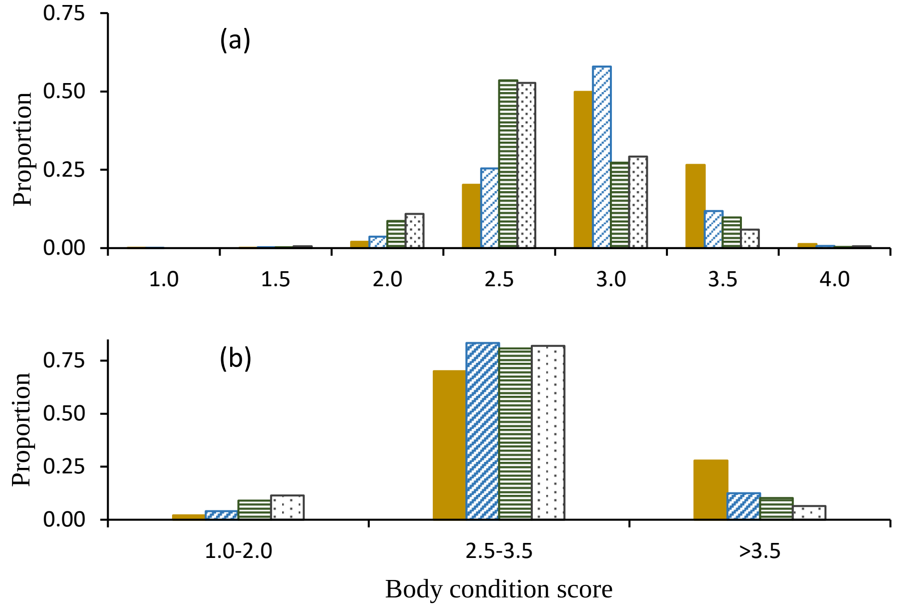


Figure 8.1 Distribution of ewe body condition scores by stage of the annual production cycle from 18,354 individual records of 5761 ewes during their fourth year (43−54 months) of age. Bar colours (grey, yellow, blue and green) indicate BCS proportions at pre-breeding, pregnancy diagnosis, pre-lambing and weaning respectively. In 8.1(a), a BCS of 1.0−4.0-point scale was used and in 8.1(b), 1.0−3.0 scale (BCS 1.0−2.0: 1, 2.5−3.5: 2 and > 3.5: 3).

To improve the balance of the BCS class distribution, a new but narrower three-class BCS scale was devised (BCS 1.0−2.0: 1, 2.5−3.5: 2 and > 3.5: 3) (Figure 1b). The selection of a new scale was guided by literature, where BCS of 2.5 to 3.5 is considered to be the range for optimal performance (Kenyon et al., 2014). Below this BCS range there is reduced performance while above this range, energy is used inefficiently. In addition, the resulting classes were re-sampled through minority class over-sampling to create "synthetic" data, a method popularly known as SMOTE (Chawla et al., 2002) using the SmoteClassif function in the UBL package (Branco et al., 2016). Resampling improves the class-level distribution (balances the number of per class observations) of a categorical variable so that the assumptions of classification models can hold.

8.2.3 Variable selection and model building

The best variable combinations for prediction of BCS (1, 2 or 3) at each stage of the annual production cycle using live weight were selected through the regularization and variable selection technique utilizing the elasticnet method in the glmnet extension (Friedman et al., 2010) in the caret package (Kuhn, 2008). The elasticnet method combines the power of two penalized-regularization methods (ridge and lasso regression) to search for significant predictors and handling of collinearity (Archer and Williams, 2012).

All models were fitted and validated using four steps as described in Chapter 6. The steps included: i) data partitioning, ii) resampling, iii) model training and iv) validation. Data were partitioned with stratification into training and testing datasets in a ratio of 3:1, with replacement. Resampling was done using the bootstrapping and aggregation (Tropsha et al., 2003) procedures in the caret package (Kuhn, 2008). During resampling, 10-equal sized sub-samples, repeated three times were selected from the dataset. Prediction models were trained on nine sub-sample sets which were used to compute the parameters and the 10th was used to evaluate the model as well as compute the error. The procedure was run 30 times (10-folds repeated three times) and the average parameter values and their probabilities were computed and computed through bootstrapping as described in Chapter 6.

The algorithms used for this work were selected from a range of probabilistic and non-probabilistic methods in order to cover the most commonly used machine learning algorithms (Valletta et al., 2017; Khaledian and Miller, 2020). A summary of the concepts, advantages and disadvantages of each algorithm is given in the appendix Table 8.1. Further, the criteria for selecting these methods included (i) successfully application in other animal science studies (Shahinfar and Kahn, 2018; Shahinfar et al., 2019; Bakoev et al., 2020), and (ii) ability to handle multi-class categorization (Leevy et al., 2018). Three traditional (ordinal logistic, multinomial regression (Agresti and Kateri, 2011; Torgo, 2016) and Linear Discriminant Analysis (LDA) (Zhao et al., 2018) statistical models (white box or low-level machine learning models), two low-level black models (Random Forest (RF) (Rennie et al., 2003) and classification and regression trees (CART) (Zhu et al., 2018) and four high-level black box models (Support Vector Machines (SVM) (Zeng et al., 2008) and K-Nearest Neighbours (KNN) (Breiman, 1998; Sun and Huang, 2010), Artificial Neural Networks (ANN), and Gradient boosting decision trees (XGB) (Ebrahimi et al., 2019) were compared. Machine learning models can be categorized in two main ways namely: (i) whether data provides labels that classify variables (supervised) or not (unsupervised) (Fisher, 1987), (ii) if a clear description of the analysis detailing how covariates and the target variable are related (classical statistical methods or white boxes), a partial description blue print (low-level- or semi-black boxes) or no description can be given (high-level black boxes) (Khaledian and Miller, 2020). All algorithms were implemented in R package using several caret package extensions (nnet, multinom, polr, lda, rpart, svmLinear, xgblinear, rf and knn (http://topepo.github.io/caret/index.html). A chart summarizing the model building and evaluation procedures is given as in Appendix IX Figure 1.

Table 8.1 Key model performance characteristics of common machine learning algorithms (selecting the most appropriate algorithms).

Model1	Concept[2]	Parameter and processes required[3]	Sample size and data dimensionality	Assumptions and data requirements	Covariate pools[4]	Computational time	Interpretability[5]	Prone to overfitting	References
Ordinal	Probabilistic regression	No hyperparameters	Affected by small sample sizes	proportional odds, linearity	No	Fast	White box	Yes	(Liao, 1994; Agresti, 1999; Agresti and Kateri, 2011)
Multinom	Probabilistic regression	No hyperparameters	Yes	proportional odds, linearity	No	Fast	White box	Yes	(Böhning, 1992; Liao, 1994; Agresti and Kateri, 2011)
LDA	Dimension reduction + separability between classes	No hyperparameters	Affected by small sample sizes, Good for high dimension data	Normality, linearity & continuous independent variables	No	Fast	White box	Yes	(Chen et al., 2000; Yu and Yang, 2001; Zheng et al., 2004)
CART	Decision trees and regression	Hyperparameters	Performs well with large datasets	numerical or categorical outcome	can remove redundant covariates	Fast	Low-level black box	No	(Quinlan, 1986; Quinlan, 1987)
RF	Decision trees, regression and bugging	Up to three hyperparameters	Performs well on small & high dimensionality data	numerical or categorical outcome	can remove redundant covariates	Decreases with sample size	Low-level black box	No	(Ho, 1995; Khaledian and Miller, 2020)
XGB	Regression trees + gradient boosting	Hyperparameter	Require large datasets	numerical or categorical outcome	can remove redundant covariates	Very fast	High-level black box	Yes, if large number of trees	(Chen and Guestrin, 2016; Zhang and Zhan, 2017)
KNN	Regression curve + hyperparameter (k)	One hyperparameter	Not good for large & high dimensionality data	No assumptions but requires scaled data	No	Decreases with sample size	Fairly interpretable	Yes	(Imandoust and Bolandraftar, 2013; Khaledian and Miller, 2020)

SVM	Maximal margins + kennel functions	Two hyperparameters	Not good for high dimension data	No assumptions	No	Decreases with sample size	High-level black box	Yes	(Gunn, 1998; Durgesh and Lekha, 2010)
ANN	Nodes (artificial neurons)	Up to seven hyperparameters	Sensitive to sample size and data dimensionality	numerical or categorical outcome	No	computationally very expensive and time consuming	High-level black box	Yes	(Tu, 1996; Daniel, 2013)

[1]Model (Ordinal: Ordinal logistic regression, Multinorm: Multinomial regression, LDA: Linear Discriminant Analysis, CART: Classification and regression tree, RF: Random Forest, XGB: Gradient boosting decision trees model, KNN: K-Nearest Neighbours, SVM: Support Vector Machines, ANN: Neural Networks). [2]Concept: How the algorithm works. [3]Parameter and processes: Tuning parameters for the algorithm. [4]Covariate pools: Intrinsic ability to remove redundant variables or to select important variables. [5]Interpretability: White box: clear model structure with parameters: black boxes: model structure and the relationship between variables is unknown.

NB: The criteria used to summarize the key model performance characteristic was a modified version of 5-point criteria by Khaledian and Miller (2020).

8.2.4 Model performance evaluation

Using a three-class BCS scale (1.0–2.0, 2.5–3.5, >3.5), model fit and ranking between models were assessed using overall accuracy, balanced accuracy, precision, F-measure, sensitivity and specificity. The metrics were computed from the number of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions as described by Tharwat (2020). In addition, Cohen's kappa statistic (Cohen, 1960), a common measure to calculate agreement between the classification of qualitative observations, was calculated as described by McHugh (2012) and Botchkarev (2019). To evaluate the power of the algorithms to correctly classify ewe BCS, measures of the balance (authenticity and prediction power) between sensitivity and specificity were computed. These indicators of model power and authenticity (Positive likelihood ratio: PLR, Negative likelihood ratio: PLN and Youden's index) combine sensitivity and specificity to emphasize how good a model can predict the outcome (Lan et al., 2017). A detailed description of the metrics (accuracy and authenticity) used in model assessment is given in Table 8.2.

Table 8.2 Model performance evaluation metrics.

Model	Definition	Formula
Balanced accuracy	The proportion of correctly classified subjects for each class. Useful especially when there is class imbalance.	$Accuracy = \frac{\frac{TP}{TP+FN} + \frac{TN}{TN+FP}}{2}$
Precision	The proportion of correctly classified subjects for a given class given that they truly belonged to that class	$Precision = \frac{TP}{TP+FP}$
F-measure	The harmonic mean of the precision and sensitivity best if there is some sort of balance between precision & sensitivity.	$F - measure = \frac{2 * (sensitivity * precision)}{sensitivity + precision}$
Sensitivity	The proportion of correctly classified subjects for a given class to those who truly belong to that class.	$Sensitivity = \frac{TP}{TP+FN}$
Specificity	The proportion of subjects correctly classified as not belonging to a given class to those that truly do not belong to that class.	$Specificity = \frac{TN}{TN+FP}$
Positive likelihood ratio (PLR)	is the ratio between the true positive and the false positive rates for "positive" events that are detected by a model.	$PLR = \frac{Sensitivity}{1 - Specificity}$
Negative likelihood ratio (NLR)	is the ratio between the false negative and true negative rates and mirrors the probability for "negative" events to be detected by a model.	$NLR = \frac{1 - Sensitivity}{Specificity}$

Youden's index (YI)	is the sum of sensitivity and specificity minus one	$YI = (Sensitivity + Specificity) - 1$
Cohen's kappa (κ)	Measures the degree of agreement between two raters or ratings (inter-rater or interrater reliability)	$\kappa = \frac{p_o - p_e}{1 - p_e}$

where: TP = true positive, TN = true negative, FP = false positive, FN = false negative,
κ = Cohen's kappa statistic, p_o = actual observed agreement, and p_e represents chance agreement.

The analysis generated a dataset of 108 records (4 time points, 3 BCS classes and 9 models) of two groups of model performance evaluation metrics firstly, the indicators of accuracy: balance accuracy, precision and F-measure, and secondly measures of model authenticity: sensitivity and specificity). To obtaina holistic picture of the overall model performance, the two groups of performance metrics were examined. Initially, each group of variables was explored using principal component analysis (PCA) to determine the appropriate number of components of dimensions where the Eigen values associated with each component were compared with those generated through a probabilistic process based on Monte Carol PCA for parallel analysis simulation (Horn, 1965; Glorfeld, 1995). Monte Carlo PCA simulated Eigen values allow comparisons based on the same sample size and number of variables. If the Eigen value of a component from real data is greater than the simulated one, then that component is important. Otherwise, if equal to or less, such components are considered not important. Consequently, one component was considered important from each group of variables (Indicators of accuracy: explained variance = 87%, indicators of Sensitivity-Specificity: explained variance = 61%) having explained most of the variability in the group data.

Principal Component Analysis is limited to continuous data. In order to decipher the patterns in the relationship between the categorical variable (BCS) and each model, regarding their overall performance, a correspondence analysis was required. Therefore, the FAMD function in the FactoMiner package (Lê et al., 2008) was used to analyze both groups of variables. The FAMD extension combines PCA and multiple correspondence analysis (MCA) to conduct factor analysis. Each group of variables then resulted in a single dimension (latent variable). A scatterplot of Accuracy and Sensitivity-Specificity latent variables was constructed for each of the four stages of the annual sheep weighing cycle. Models were ranked on a scale of 1 to 9 (where 1 is best and 9 the poorest) at each stage of the annual production cycle, to obtain the overall performance rank.

8.3 Results

8.3.1 Overall performance of machine learning models

This section presents results for the accuracy in a broad sense, sensitivity, and specificity of nine models in predicting ewe BCS based on the testing dataset (Tables 8.3 and 8.4).

Additionally, Appendix IX Table 1 gives the model accuracy comparisons between across stages of the annual sheep weighing cycle in New Zealand.

Results showed that there were significant ($p < 0.05$) differences in model prediction performance based on the Boniferroni p-value adjustment method for pairwise comparisons (Appendix IX Table 2). The gradient boosting decision tree algorithm (XGB) had the highest ($p < 0.05$) accuracy (average = 90.3%) and kappa statistic (κ = 82.1%) at pre-breeding, pregnancy diagnosis, pre-lambing and weaning, making it the most accurate algorithm for ewe BCS prediction on the 1 to 3 (1.0−2.0; 2.5−3.5; >3.5) scale (Table 8.3). The RF (Appendix IX: Table 2, Figure 2) algorithm had a slightly lower but good accuracy making it the best alternative to XGB. The Multinorm, LDA, Ordinal and CART algorithms had moderate to fair accuracies. Pre-lambing, XGB and RF were comparable and had the highest accuracies. The Random Forest and K-Nearest Neighbours (KNN) in decreasing order were also considered good prediction models having scored above 80% accuracy and 70% kappa statistics at all times of the year. The CART algorithm consistently gave the lowest ($p > 0.05$) accuracy except pre-lambing where its accuracy was comparable ($p = 0.047$; Appendix D2) to that of ordinal logistic regression. The lowest average accuracy was 66.6% seen for the CART model at weaning (Table 8.3, parenthesis). Overall, all algorithms had greater accuracy than a random guess (i.e. accuracy = 33.3%) in classifying BCS.

In terms of overall authenticity, models were biased towards being more specific than sensitive (Table 8.4). The ranking of model authenticity followed a trend like that of accuracy. The gradient boosting decision tree algorithm (XGB) had the highest sensitivity (average = 87.7%) as well as specificity (average = 93.9%) across all stages of the annual sheep weighing cycle, making it the most authentic and powerful algorithm for categorizing ewe into the correct BCS classes on 3-point scale (1.0−2.0; 2.5−3.5; >3.5) (Table 8.3). The XGB model was closely followed by RF (average sensitivity = 85.5%, average specificity: 92.8%) while CART (average sensitivity: 58.7%, average specificity: 79.5%) was the poorest.

Table 8.3 Accuracy and Kappa statistics of nine predictive models for ewe BCS at 43−54 months of age at different stages of the annual production cycle. Values in parenthesis denote the minimum and maximum accuracy, in ascending order. The superscripts ^{123} where 1: 1.0−2.0, 2: 2.5−3.5 and 3: >3.5 indicate the BCS class from which the value was observed. The first superscript indicates the class from which the minimum estimate was observed, while the second value indicates the class from which the maximum estimate was achieved). All models were significant ($p < 0.05$) and better than a random guess (i.e. Accuracy = 33.3%). All ewe BCS predictions were based on current and previous live weight.

	Pre-breeding		Pregnancy diagnosis		Pre-lambing		Weaning	
Model	Accuracy	Kappa(κ)	Accuracy	Kappa(κ)	Accuracy	Kappa(κ)	Accuracy	Kappa(κ)
XGB	89.5 (85.6−97.5)^{3,1}	79.6	91.2 (88.5−93.4)^{3,1}	82.3	90.6 (88.8−91.4)^{2,1}	82.9	91.7 (90.1−93.2)^{1,3}	83.4
RF	89.0 (84.7−96.6)^{2,1}	78.0	90.0 (87.5−92.9)^{3,1}	78.0	89.2 (86.6−91.6)^{2,3}	78.5	88.6 (88.2−89.6)^{1,3}	77.1
KNN	87.0 (81.2−95.7)^{2,1}	75.5	86.8 (84.7−89.8)^{3,1}	75.5	86.2 (83.0−89.7)^{2,3}	66.0	86.4 (84.6−88.8)^{2,3}	77.7
SVM	86.7 (78.8−96.6)^{2,1}	75.9	88.5 (84.8−93.1)^{2,1}	73.7	73.8 (72.0−74.7)^{2,1}	71.7	88.8 (85.3−91.2)^{2,3}	72.7
ANN	85.2 (79.0−94.2)^{2,1}	72.2	82.0 (80.5−85.1)^{2,1}	65.6	78.9 (75.5−82.4)^{1,3}	69.5	84.0 (82.0−86.9)^{1,3}	68.0
Multinorm	82.7 (76.4−91.7)^{2,1}	66.8	77.6 (73.8−80.0)^{3,1}	56.1	73.5 (71.8−75.1)^{1,3}	48.8	75.9 (74.4−78.1)^{3,2}	51.8
LDA	81.2 (73.8−91.1)^{2,1}	63.6	77.1 (72.2−79.6)^{3,1}	54.6	73.8 (71.5−75.5)^{1,3}	49.5	75.9 (74.4−78.7)^{1,2}	51.7
Ordinal	79.6 (70.7−88.4)^{2,1}	48.4	72.7 (67.6−75.8)^{2,1}	47.7	68.4 (58.7−74.8)^{2,3}	37.0	72.4 (67.8−76.2)^{2,1}	44.9
CART	72.6 (58.6−85.1)^{2,1}	47.3	69.8 (64.0−73.3)^{3,1}	40.5	67.5 (62.8−71.1)^{1,2}	41.8	66.6 (61.4−70.1)^{2,1}	33.2

Model: (XGB: Gradient boosting decision trees model, RF: Random Forest, KNN: K-Nearest Neighbours, SVM: Support Vector Machines, ANN: Neural Networks, Multinorm: Multinomial regression, LDA: Linear Discriminant Analysis, Ordinal: Ordinal logistic regression, CART: Classification and regression tree).

Table 8.4 Indicators of authenticity (sensitivity and specificity) of nine predictive models for ewe BCS at 43−54 months of age at different stages of the annual production cycle. Values in parenthesis denote the minimum and maximum sensitivity or specificity, in ascending order. The superscripts ^{123} where 1: 1.0−2.0, 2: 2.5−3.5 and 3: >3.5 indicate the BCS class from which the value was observed. In their sequence, the first superscript indicates the class from which the minimum estimate was observed, while the second value indicates the class from which the maximum estimate was achieved). All ewe BCS predictions were based on current and previous live weight.

	Pre-breeding		Pregnancy diagnosis		Pre-lambing		Weaning	
Model	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
XGB	86.0 (79.7–96.3)^{3,1}	93.1 (89.1–98.9)^{2,1}	88.2 (83.7–90.4)^{3,1}	94.2 (93.1–96.3)^{2,1}	87.5 (85.9–88.8)^{1,3}	93.8 (89.7–97.5)^{2,1}	89.0 (84.8–92.3)^{1,2}	94.5 (91.6–96.5)^{2,3}
RF	85.3 (80.0–95.3)^{2,1}	92.8 (89.3–97.9)^{2,1}	86.7 (80.9–90.3)^{3,1}	93.4 (90.5–95.5)^{2,1}	85.6 (82.6–88.6)^{1,3}	92.8 (87.5–96.4)^{2,1}	84.8 (82.5–87.6)^{1,2}	92.4 (88.9–93.4)^{2,3}
SVM	82.6 (74.8–93.8)^{2,1}	91.4 (87.5–97.5)^{2,3}	82.3 (75.3–84.2)^{3,2}	91.2 (84.2–95.4)^{2,1}	81.5 (73.5–86.1)^{1,3}	90.8 (81.1–98.1)^{2,1}	81.9 (77.6–85.6)^{3,2}	90.9 (83.5–95.1)^{2,3}
KNN	82.2 (66.8–96.2)^{2,1}	91.2 (85.9–97.0)^{3,1}	84.7 (75.5–91.8)^{2,1}	92.3 (88.4–94.5)^{3,1}	65.0 (63.0–67.3)^{1,2}	82.5 (76.8–86.4)^{2,1}	85.1 (78.6–88.9)^{2,3}	92.6 (91.9–93.6)^{2,3}
ANN	80.2 (71.3–91.7)^{2,1}	90.2 (86.7–96.7)^{2,1}	76.0 (73.2–78.0)^{3,1}	88.0 (84.3–92.2)^{2,1}	71.8 (56.5–80.2)^{1,3}	85.9 (78.8–94.4)^{2,1}	78.7 (70.5–84.1)^{1,2}	89.3 (82.4–93.5)^{2,1}
Multinom	76.8 (68.5–89.0)^{2,1}	88.5 (84.4–94.5)^{2,1}	70.0 (62.7–71.4)^{3,2}	85.1 (81.8–88.7)^{2,1}	64.7 (58.6–68.7)^{1,3}	82.4 (80.6–84.9)^{2,1}	67.9 (63.3–76.2)^{3,1}	83.9 (80.1–86.2)^{2,1}
LDA	74.9 (64.7–87.7)^{2,1}	87.6 (82.8–94.4)^{2,1}	69.4 (57.1–82.7)^{3,2}	84.8 (76.6–90.7)^{2,1}	65.0 (56.3–69.4)^{1,3}	82.5 (79.2–86.8)^{2,1}	67.8 (61.5–79.8)^{3,2}	83.9 (77.6–87.4)^{2,3}
Ordinal	72.7 (61.6–82.4)^{2,1}	86.5 (79.7–94.5)^{2,1}	63.6 (60.7–67.9)^{2,3}	81.7 (73.1–90.9)^{2,1}	57.9 (41.4–69.3)^{2,3}	79.0 (76.1–80.8)^{2,1}	63.2 (58.3–68.5)^{3,1}	81.6 (72.8–88.2)^{2,3}
CART	63.3 (37.0–82.5)^{2,1}	81.9 (77.6–87.8)^{3,1}	59.7 (41.1–77.3)^{3,2}	80.0 (67.1–86.0)^{2,3}	56.7 (37.9–72.3)^{1,2}	78.3 (71.2–87.7)^{2,1}	55.4 (39.2–62.9)^{2,1}	77.7 (72.4–83.6)^{3,2}

Model: (XGB: Gradient boosting decision trees model, RF: Random Forest, KNN: K-Nearest Neighbours, SVM: Support Vector Machines, ANN: Neural Networks, Multinorm: Multinomial regression, LDA: Linear Discriminant Analysis, Ordinal: Ordinal logistic regression, CART: Classification and regression tree).

In the following section we present results for the construct or latent variables which are representative of the three specific measures of model accuracy (class-level or balanced Accuracy, Precision and F-measure) together with two indicators of predictive power/authenticity (Sensitivity, Specificity) across four stages of the annual sheep weighing cycles (Figures 8.2, 8.3, 8.4 and 8.5). A summary of the indicators of accuracy and authenticity was provided in Tables 8.3 and 8.4. Additionally, Appendix D4 provides extra two measures of accuracy (Precision and F-measure) used in the construction of the accuracy latent variable. The results show the patterns in the relationship between the latent variables with BCS class prediction for each model. The CART model having had the lowest accuracy and power measures across all stages of the annual sheep weighing cycle and was selected as the reference for comparisons.

8.3.1.1 Pre-breeding

At pre-breeding, the models had a clear-cut hierarchy in performance, with XGB being the best and CART the poorest (Figure 2). The XGB was the best algorithm with 17% more accuracy than CART, which was the least accurate in predicting ewe BCS (Table 8.3). The best balance between accuracy and authenticity (points along or touching the diagonal line) was observed in the moderate performing models including ANN, Multinom, LDA and Ordinal (Figure 8.3). The best performing models (XGB, RF, SVM and KNN) were biased towards accuracy while the poorest (CART) was biased towards authenticity. In terms of BCS, the best accuracy was achieved in the 1.0−2.0 class and the lowest in the 2.5−3.5 class for all models except for XGB which was least accurate in the >3.5 class. The best accuracy (97.5%) was achieved using the XGB in the 1.0−2.0 BCS class and the lowest (58.6%) was observed using the CART algorithm in the 2.5−3.5 class (Table 8.3, parenthesis).

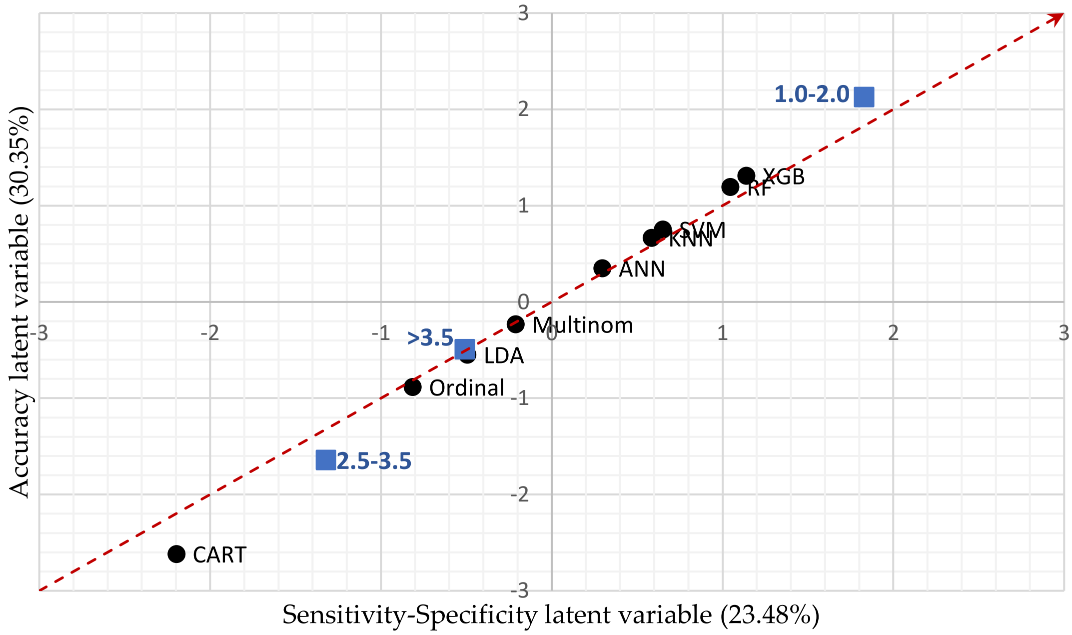


Figure 8.2 A plot of the accuracy and sensitivity-specificity latent variables from their first dimension/component obtained through a factor analysis of mixed variables (a combination of Principal Component and Multiple Correspondence Analyses) procedure on measures of performance for the prediction of ewe BCS during pre-breeding. Dots (red sphere: model, blue square: BCS class). Dotted diagonal line indicates a balance between accuracy and sensitivity-specificity. If dot is above, then model or BCS class was more accurate than sensitive-specific while the reverse indicates that the model was more sensitive than accurate. The further and more positive a model is along the diagonal line, the greater and better is its prediction power. The variance explained by each extracted first dimension for each latent variable (Accuracy, sensitivity-specificity) is given in parenthesis along the axes.

All models were most sensitive to the 1.0−2.0 class and least sensitive to the 2.5−3.5 class except XGB which was least sensitive to the >3.5 class. The XGB was the best algorithm being 23% more sensitive than CART, which was the least sensitive in predicting ewe BCS (Table 8.3). The highest BCS classification sensitivity was observed using XGB and KNN models (96.3%) in the 1.0−2.0 BCS class while CART (37.0%) had the lowest in the 2.5−3.5 class (Table 8.4, parenthesis). All models had the highest specificity observed in the 1.0−2.0 BCS class except for SVM which had the highest specificity in the >3.5 class and both KNN and CART which had their lowest in the >3.5 class. The XGB was the best algorithm with 12% more specificity than CART, which had the least specificity in predicting ewe BCS (Table 8.4). The highest specificity (98.9%) was observed in the 1.0−2.0 class for XGB and the lowest (72.6%) in the >3.5 class for CART model (Table 8.4, parenthesis).

8.3.1.2 Pregnancy diagnosis

At pregnancy diagnosis, the models had a clear-cut hierarchy in performance, with XGB being the best and CART the poorest (Figure 8.3). The Multinom and LDA models were closely juxtaposed indicating that they had comparable performance. The XGB was the best algorithm with 21% more accuracy than CART, which was the least accurate in predicting ewe BCS (Table 8.3). The best balance between accuracy and authenticity was observed in the ANN model. The XGB, RF, SVM and KNN models were biased towards accuracy while the Multinom, LDA, Ordinal and CART were biased towards authenticity (Figure 8.3). In terms of BCS, the best accuracy was achieved in the 1.0−2.0 class and the lowest in the >3.5 class for all models except for SVM, ANN and Ordinal which were least accurate in the 2.5−3.5 class. The highest accuracy (93.4%) was achieved using the XGB in the 1.0−2.0 BCS class and the lowest (64.0%) was observed using the CART algorithm in either the >3.5 class (Table 8.3, parenthesis).

There was no clear pattern in class-level model sensitivity at pregnancy diagnosis. The XGB was the best algorithm with 29% more sensitivity than CART, which was the least sensitive in predicting ewe BCS (Table 8.4). The highest BCS classification sensitivity was observed using KNN models (91.8%) in the 1.0−2.0 BCS class while CART (41.1%) had the lowest in the >3.5 class (Table 8.3, parenthesis). All models had the highest specificity observed in the 1.0−2.0 BCS class except for CART which had the its highest in the >3.5 class. The XGB was the best algorithm with 14% more specificity than CART, which had the least specificity in predicting ewe BCS (Table 8.4). The highest specificity (96.3%) was observed in the 1.0−2.0 class for XGB and the lowest (67.1%) in the 2.5−3.5 class for CART model.

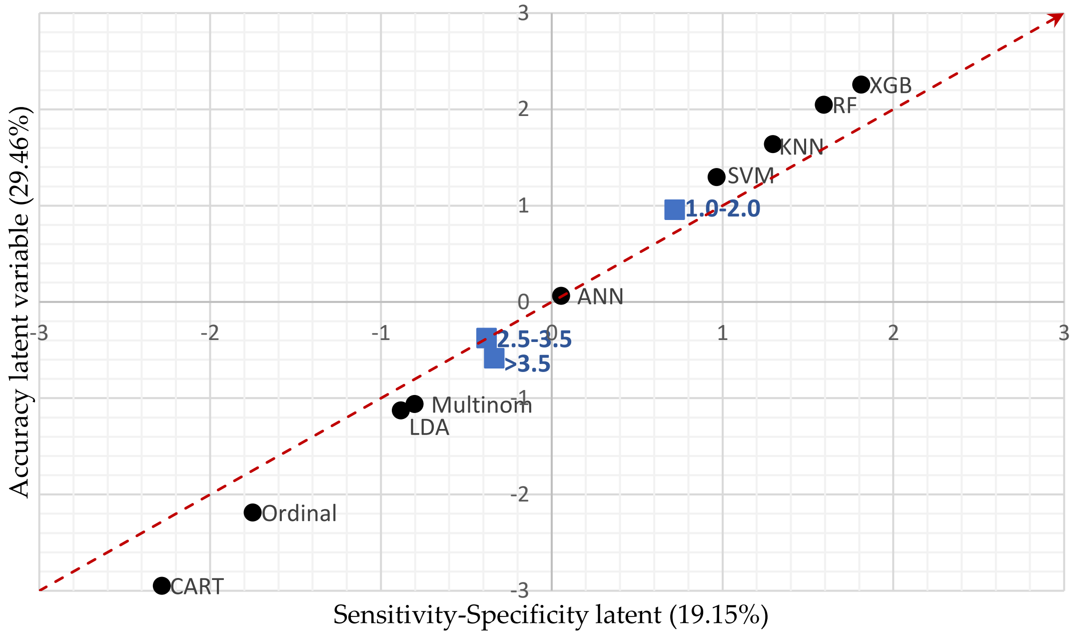


Figure 8.3 A plot of the accuracy and sensitivity-specificity latent variables from their first dimension/component obtained through a factor analysis of mixed variables (a combination of Principal Component and Multiple Correspondence Analyses) procedure on measures of performance for the prediction of ewe BCS during pregnancy diagnosis. Dots (red sphere: model, blue square: BCS class). Dotted diagonal line indicates a balance between accuracy and sensitivity-specificity. If dot is above, then model or BCS class was more accurate than sensitive-specific while the reverse indicates that the model was more sensitive than accurate. The further and more positive a model is along the diagonal line, the greater and better is its prediction power. The variance explained by each extracted first dimension for each latent variable (Accuracy, sensitivity-specificity) is given in parenthesis along the axes.

8.3.1.3 Pre-lambing

At pre-lambing, the models had a clear-cut hierarchy in performance, with XGB being the best and CART the poorest (Figure 8.3). It was worth noting that, the KNN model which had been among the best four models at pre-breeding and pregnancy diagnosis, was downgraded into a moderate model. The KNN, Multinom and LDA models had overlapping overall performance. The XGB was the best algorithm with 23% more accuracy than CART, which was the least accurate in predicting ewe BCS (Table 8.3). The best overall accuracy was achieved in the >3.5 BCS class and the lowest in the 2.5−3.5 class (Table 8.3, parenthesis). Regarding BCS class-level model accuracy, there was no clear pattern. The majority of the models (RF, KNN, ANN, Multinom, LDA and Ordinal) were most accurate in the >3.5 BCS class and least accurate in the 2.5−3.5 class. The least accuracy for majority of the models (XGB, RF, KNN, SVM and Ordinal) was observed in the 2.5−3.5 class. The highest accuracy (92%) was achieved using the RF model in the >3.5 BCS class and the lowest (63 %) was observed using the CART algorithm in either the 1.0−2.0 class (Table 8.3, parenthesis).

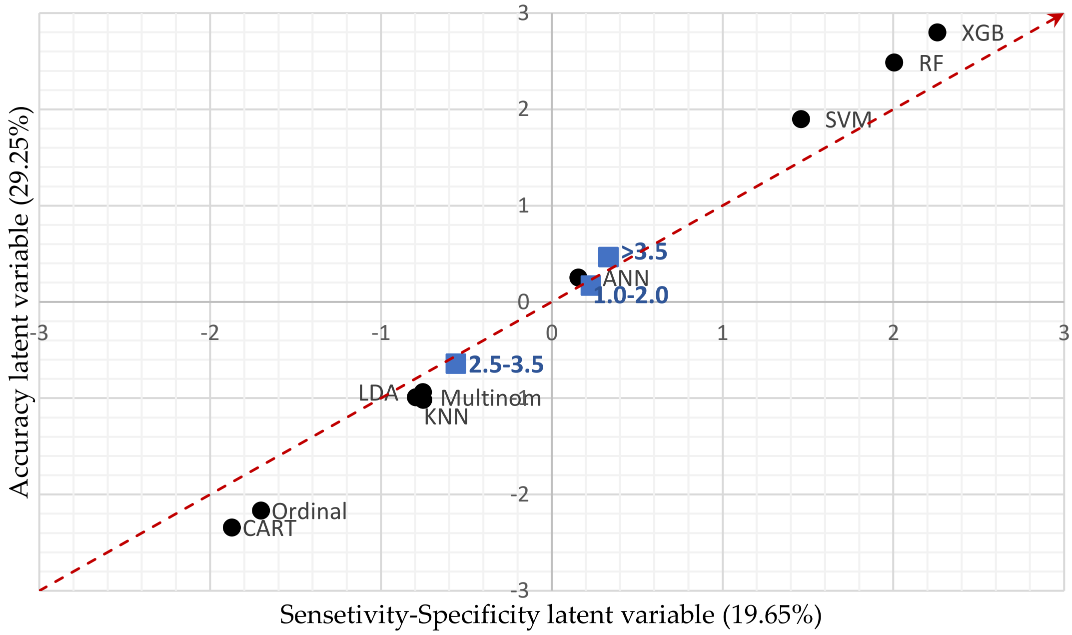


Figure 8.4 A plot of the accuracy and sensitivity-specificity latent variables from their first dimension/component obtained through a factor analysis of mixed variables (a combination of Principal Component and Multiple Correspondence Analyses) procedure on measures of performance for the prediction of ewe BCS at pre-lambing. Dots (red sphere: model, blue square: BCS class). Dotted diagonal line indicates a balance between accuracy and sensitivity-specificity. If dot is above, then model or BCS class was more accurate than sensitive-specific while the reverse indicates that the model was more sensitive than accurate. The further and more positive a model is along the diagonal line, the greater and better is its prediction power. The variance explained by each extracted first dimension for each latent variable (Accuracy, sensitivity-specificity) is given in parenthesis along the axes.

All models were biased with XGB, RF, SVM and ANN inclined towards accuracy, while KNN, Multinon, LDA, Ordinal and CART were inclined towards authenticity (Figure 8.4). Further, all models were most sensitive to the >3.5 class and least sensitive to the 1.0−2.0 class except KNN and CART with the highest sensitivity in the 2.5−3.5 class and Ordinal with the lowest sensitivity in the 2.5−3.5 class. The XGB was the best algorithm with 31% more sensitive than CART, which was the least sensitive in predicting ewe BCS (Table 8.4). The highest BCS classification sensitivity was observed using XGB models (88.8%) in the >3.5 BCS class while CHART (37.9%) had the lowest in the 1.0−2.0 class (Table 8.4, parenthesis). All models had the highest specificity observed in the 1.0−2.0 BCS class. The XGB was the best algorithm with 16% more specificity than CART, which had the least specificity in predicting ewe BCS. The highest specificity (97.5%) was observed in the 1.0−2.0 class for XGB and the lowest (71.2%) in the 2.5−3.5 class for CART model (Table 8.4, parenthesis).

8.3.1.4 Weaning

At weaning, the models had a clear-cut hierarchy in performance, with XGB being the best and CART the poorest (Table 8.3; Figure 8.5). The RF and KNN models had overlapping overall performance. The XGB was the best algorithm with 33% more accuracy than CART, which was the least accurate in predicting ewe BCS (Table 8.3). The majority of the models were biased towards accuracy, except for Multinon, LDA, Ordinal and CART, which were inclined towards authenticity (Figure 8.5). The best overall accuracy was achieved in the >3.5 BCS class and the lowest in the 2.5−3.5 class. Regarding the BCS level model accuracy, there was no clear pattern. However, majority of the models (XGB, RF, SVM, KNN and ANN) were most accurate in the >3.5 BCS class. The least model accuracy was equally observed in the 1.0−2.0 and 2.5−3.5 BCS classes, across models. The highest accuracy (93.2%) was achieved using the RF model in the >3.5 BCS class and the lowest (61.4 %) was observed using the CART algorithm in either the 2.5−3.5 class (Table 8.3, parenthesis).

There was no clear pattern in class-level model sensitivity at weaning. The XGB was the best algorithm with 34% more sensitivity than CART, which was the least sensitive in predicting ewe BCS (Table 8.4). The highest BCS classification sensitivity was observed using XGB models (92.3%) in the 2.5−3.5 BCS class while CHART (39.2%) had the lowest in the 2.5−3.5 class (Table 8.4, parenthesis). All models had the highest specificity observed in the >3.5 BCS class and the least in the 2.5−3.5 class, except for the CART whose specificity arrangement was the opposite and for ANN and Multinom which had their highest specificity in the 1.0−2.0 class. The XGB was the best algorithm with 17% more specificity than CART, which had the least specificity in predicting ewe BCS (Table 8.4). The highest specificity (96.5%) was observed in the 1.0−2.0 class for XGB and the lowest (72.4%) in the 2.5−3.5 class for CART model (Table 8.4, parenthesis).

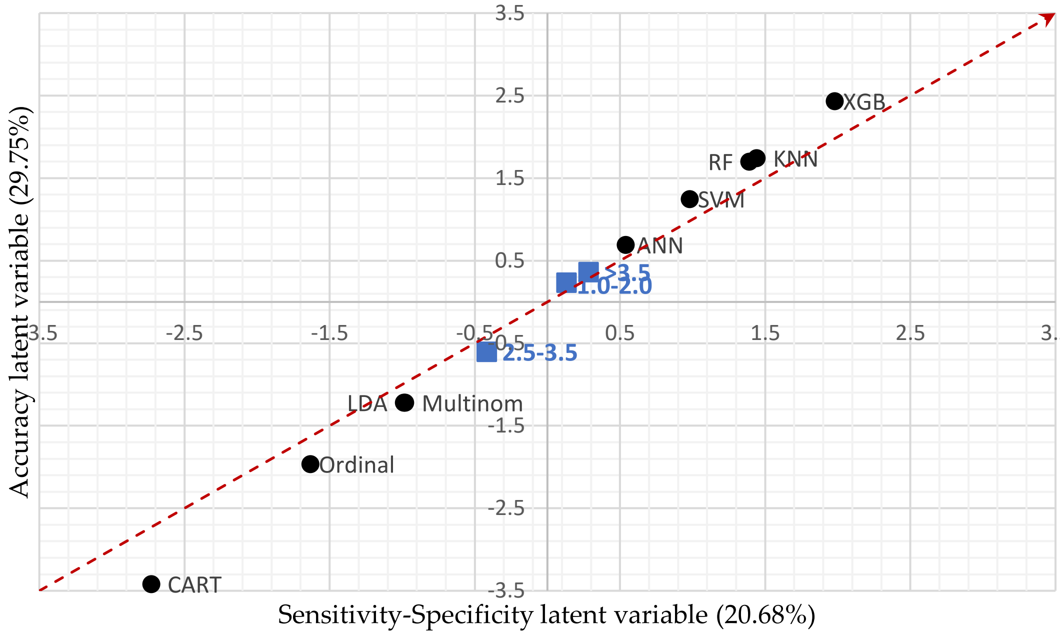


Figure 8.5 A plot of the accuracy and sensitivity-specificity latent variables from their first dimension/component obtained through a factor analysis of mixed variables (a combination of Principle Component and Multiple Correspondence Analyses) procedure on measures of performance for the prediction of ewe BCS at weaning. Dots (red sphere: model, blue square: BCS class). A plot of the accuracy and sensitivity-specificity latent variables from the first dimension/component obtained through a factor analysis of mixed variables (a combination of Principal Component Analysis and Multiple Correspondence Analysis) procedure on measures of performance for the prediction of ewe BCS at weaning. Dots (red sphere: model, blue square: BCS class). Dotted diagonal line indicates a balance between accuracy and sensitivity-specificity. If dot is above, then model or BCS class was more accurate than sensitive-specific while the reverse indicates that the model was more sensitive than accurate. The further and more positive a model is along the diagonal line, the greater and better is its prediction power. The variance explained by each extracted first dimension for each latent variable (Accuracy, sensitivity-specificity) is given in parenthesis along the axes.

8.3.2 The balance between sensitivity and specificity

The data showed that the overall specificity 86% (67−98%) was higher than sensitivity 74% (37−96%) values across all algorithms (Table 3). An assessment of the indicators of the balance between sensitivity and specificity was undertaken and the indices are summarized in Table 4. The positive likelihood ratio (PLR) for all models were greater than 1.0 while the negative likelihood ratio (NLR) were less than 1.0 across stages of the annual production cycle. The XGB model had the highest PLR and lowest NLR while CART had the lowest PLR and highest NLR across stage of the annal cycle. Similarly, Youden’s index, YI was consistently highest for XGB model and lowest for the CART model.

Table 8.5 Measures of the balance between sensitivity and specificity of the BCS prediction models by stage of the annual production cycle.

	Pre-breeding			Pregnancy diagnosis			Pre-lambing			Weaning		
Model	PLR	NLR	YI	PLR	NLR	YI	PLR	NLR	YI	PLR	NLR	YI
XGB	33.41	0.15	0.79	16.48	0.13	0.82	19.39	0.13	0.81	18.32	0.12	0.83
RF	20.49	0.16	0.78	14.45	0.14	0.80	15.33	0.16	0.78	12.25	0.16	0.77
SVM	16.88	0.19	0.74	12.13	0.19	0.74	18.48	0.20	0.72	11.79	0.20	0.73
KNN	15.21	0.20	0.73	12.3	0.17	0.77	3.90	0.42	0.48	11.64	0.16	0.78
ANN	13.04	0.22	0.70	6.94	0.27	0.64	6.32	0.32	0.58	8.66	0.24	0.68
Multinom	8.65	0.27	0.65	4.87	0.35	0.55	3.69	0.43	0.47	4.28	0.38	0.52
LDA	8.16	0.29	0.62	5.12	0.36	0.54	3.78	0.42	0.48	4.37	0.38	0.52
Ordinal	7.66	0.32	0.59	4.20	0.45	0.45	2.83	0.54	0.37	3.83	0.45	0.45
CART	3.92	0.46	0.45	3.27	0.49	0.40	2.70	0.54	0.35	2.49	0.57	0.33

Models: (XGB: Gradient boosting decision trees model, RF: Random Forest, KNN: K-Nearest Neighbours, SVM: Support Vector Machine, ANN: Neural Networks, Multinorm: multinomial regression, LDA: Linear Discriminant Analysis, Ordinal: Ordinal logistic regression, CART: Classification and Regression Tree). Measures of the balance between sensitivity and specificity (PLR: Positive likelihood rate, NLR: Negative likelihood rate and YI: Youden's index). A good model (PLR value > 1.0 and the larger PLR is the better, NLR value less than 1.0 and the smaller the better, YI ranges from 0 to 1.0 and values that approach 1.0 show higher authenticity and prediction power).

8.3.3 Overall model ranking

Overall, black box models were better than low-level white box models (Table 5). The XGB was consistently the best performing while CART was the poorest model. There was change in model ranking across stages of the annual production cycle except for XGB, LDA, Ordinal and CART.

Table 8.5. Model ranking by stage of annual production cycle and overall.

Model	Pre breeding	Pregnancy diagnosis	Pre-lambing	Weaning	Overall
XGB	1	1	1	1	1 (1.0)
RF	3	2	2	2	2 (2.3)
SVM	4	3	4	3	3 (3.5)
KNN	2	6	3	4	4 (3.8)
ANN	5	4	5	5	5 (4.8)
Miltinom	6	5	6	6	6 (5.8)
LDA	7	7	7	7	7 (7.0)
Ordinal	8	8	8	8	8 (8.0)
CART	9	9	9	9	9 (9.0)

Overall (overall rank with means in parenthesis). The lower the rank the greater the BCS prediction performance.

8.4 Discussion

The present study utilized machine learning classification algorithms to explore the possibility of predicting BCS from current and previous live weight in mature ewes (at approximately 43−54 months of age). Body condition score was treated as a categorical variable with three levels (1.0−2.0, 2.5−3.5; >3.5). Nine of the most recognized machine learning models (XGB, ANN, RF, K-NN, SVM, Ordinal, Multinom, LDA and CART models) were applied to preprocessed datasets.

We applied a strategy to reduce the accuracy and authenticity measures into two dimensions, to generate latent variables or constructs that were plotted to give a visual summary of model performance. This technique gave a visual display (a holistic picture) of overall model performance which made it easier to decipher the patterns in the relationship between the accuracy and authenticity of models in BCS prediction. Previous studies have suggested the use of several metrics to give an indication about a model's accuracy and authenticity(Hossin and Sulaiman, 2015; Botchkarev, 2019; Dinga et al., 2019; Tharwat, 2020). These have, however, been piecemeal with no unifying interface. By bringing together both accuracy and authenticity measures in a single display, we appear to have come up with a solution. This innovation could serve as a platform for interrogating even better ways of model performance evaluation.

8.4.1 Overall accuracy

The findings suggest that ewe BCS prediction from current and previous live weight can be achieved using machine learning classification algorithms within the limited BCS range used in the present study. The results indicated that XGB was the most efficient and robust model (overall accuracy = 87.6%; sensitivity = 87.7%; specificity = 93.9%). Other good alternatives to XGB for predicting ewe BCS were three algorithms (KNN, RF and SVM) with accuracies > 80% and kappas > 70% while the remaining four (CART, Ordinal, LDA and Multinomial) were weak algorithms (accuracies < 70%, kappas < 60%). All models performed better than a random guess with the most efficient ones giving prediction errors as low as 11% and 38%. According to Galdi and Tagliaferri (2018), a perfect classifier has a rate of 100% while a random guess would give a 33.3% error for three-level classifiers (Dietterich, 2000; Galdi and Tagliaferri, 2018). The weakest algorithms outperformed a random guess by only 8, 11, 15 and 20 %, respectively, using the current study data. Whereas accuracy measures can be interpreted arbitrarily, Cohen's Kappa statistic has been classified (Landis and Koch, 1977; McHugh, 2012) into six different categories, no agreement (values ≤ 0), none to slight (0.01–0.20), fair (0.21–0.40), moderate (0.41−0.60), substantial (0.61−0.80), and almost perfect agreement (0.81−1.00). Further, Fleiss et al. (1981), suggested that kappa values greater than 0.75 may be taken to represent excellent agreement beyond serendipity, values below 0.40 as poor agreement, and values between 0.40 and 0.75 as fair to good agreement. The findings in this study suggest that using the top performing algorithms (XGB and RF), ewe BCS can be predicted with high accuracy across four phases of the annual production cycle.

8.4.2 Class-level accuracy

Results also showed that at the accuracy-related, class-level, metrics including accuracy, precision and F-measure were highest for XGB making it the most efficient and robust model for ewe BCS prediction. Further, there appeared to be variability in all metrics across stages of the annual sheep weighing cycle and BCS class. This variation in accuracy across the stages of the annual production cycle suggests that with the exception of XGB, different models may be required to predict BCS at different stages of the annual production cycle. Similarly, different models may be required if there is need for greater accuracy in one BCS class than others. This is especially important when great accuracy is required for management decisions with far reaching consequences such as when limited resources must be allocated to only target classes. Further, results indicated that the higher- level (black box) machine learning models such as XGB and RF, were better at separating BCS into distinct classes than the lower-level (white box) models such as multinomial or ordinal logistic regression.

In the current study, the best balance between accuracy and authenticity (sensitivity-specificity) was achieved during pre-breeding compared with other stages of the annual production cycle. This observation could have been due to the "relative ease" to condition score a ewe pre-breeding than other stages of the annual production cycle (Kenyon et al., 2014). Prior to breeding most farmers enhance ewe feeding in a process known as flushing (Kenyon et al., 2008; Kenyon et al., 2011a) which likely resulted into uniform tissue (fat and muscle) distribution around the body. In addition, the weight measurements recorded pre-breeding are not confounded by the conceptus mass which is the case at pregnancy diagnosis and pre-lambing. The conceptus mass influences the ewe live weight from pregnancy through the pre-lambing stage (Kenyon et al., 2008; Kenyon et al., 2011b) which coincides with the two time-point weight measurements during those stages of the annual production cycle. Further, during lactation a ewe has its greatest nutrient requirements for energy and protein (Nicol and Brookes, 2007) and at weaning a ewe is drained by the lactation process leading to variability in fat deposition around the body and are consequently lighter. Using the same ewe population, a decreasing trend in ewe BCS as a ewe aged plateauing after 43−54 months has been reported (Chapter 6). This was attributed to a likelihood that farmers were under feeding their aging ewes at certain stages or periods of the annual production cycle. Lactation period could be one of such periods, resulting in failure to meet ewe dietary energy and protein requirements and consequently leading to thinner animals. The management conditions at pregnancy diagnosis, pre-lambing and weaning, therefore, could lead to differences in fat deposition around the body resulting into variability in BCS.

8.4.3 Class-level model authenticity

Among the indicators of model authenticity, the models had apparently greater specificity than sensitivity which could point to unbalanced distinguishing power to make predictions. An examination of three indicators of balance between sensitivity and specificity or model authenticity/power (PLR and YI) indicated that all models had values within acceptable authenticity and power (PLR > 1.0, NLR < 1.0 and YI > 1.0) across stage four stages of the annual production cycle indicating that, all models had balanced sensitivity and specificity. Results also showed that XGB had the highest PLR and YI and the lowest NLR. Combined with the results from the measures of accuracy, these results rank XBG as the most robust model for BCS prediction. Sensitivity is defined as the proportion of individuals or items who belong to a given BCS class and are correctly identified, while specificity is the proportion which do not belong to a given class and are excluded by the test. There exists an inverse relationship between sensitivity and specificity of a test or prediction model (Parikh et al., 2008; Naeger et al., 2013). If a model has high sensitivity, it is capable of detecting "real" BCS classes but it also faces losses from consuming more resources due to mandatory confirmatory tests (to rule out the false positives) or when the limited resources have to be given to only the right candidates. However, if a model has high specificity, the system benefits from a significant reduction in the consumption of resources, and time, but has a decreased capacity to detect "real" BCS classes, which can lead to failure to detect many events of importance (Lan et al., 2017). The higher specificity would not be advantageous, as failure to detect ewes inside or outside the BCS range (2.5−3.5) for optimum productivity would affect management decisions negatively. Therefore, a good model needs to achieve a balance between sensitivity and specificity (Obuchowski and Bullen, 2018).

This study suggests that ewe BCS prediction from current and previous live weight can usefully be achieved using machine learning classification algorithms within a limited BCS range used in the present study. This study used unadjusted live weight (i.e. confounded by factors such fleece length variations and conceptus mass from pregnancy to lambing) records alone to achieve accuracies up to 89% to assign BCS to one out of three classes. It is likely that if adjusted live weights were used together with other key variables that affect BCS, optimum accuracy would be achieved from these BCS prediction algorithms. In Chapter 7, it was suggested that the accuracy of BCS prediction could be improved if all key variables affecting the relationship between live weight and BCS were accounted for. If this was the case, the efficiency of the machine learning models tested could also be enhanced.

Although not directly comparable, having used different scale ranges and different measures of model performance, the best ML model (XGB) in the current study had great efficiency (based on live weight predictors alone), achieved greater than 90% accuracies and was stable (Accuracy: 86−93%) across stages of the annual production cycle. In their previous study (Chapter 7) based on linear regression models, only, moderate goodness of fit (R^2 = 50%) was achieved using more resources (both LW and BCS records). Further, the model goodness of fit and accuracy varied greatly (R^2: 28–64%) across stages of the annual production cycle, making the linear regression models less stable. Combined, therefore, this suggests that machine learning models would offer better BCS predictions than the linear regression models.

8.5 Conclusions

The results of the present study showed that ewe BCS (grouped) can be predicted with great accuracy on a narrow BCS (1.0–2.0, 2.5–3.5, >3.5) scale from a ewe's current and previous live weight using machine learning algorithms. The gradient boosting decision trees algorithm was the most efficient for ewe BCS prediction. The results of this study, therefore, support the hypothesis that BCS can be accurately predicted from a ewe's current and previous live weights. The algorithms having been trained on a large representative dataset, should be able to give accurate ewe BCS predictions. These algorithms (acquired intelligence) could be incorporated into weighing systems to easily and quickly give farmers ewe BCS without the need for hands-on burden. Future studies should investigate how to ameliorate the accuracy of BCS prediction and the possibility of individual BCS prediction on a full range (1.0−5.0).

Chapter 9. Predicting ewe body condition score using adjusted live weight for conceptus and fleece weight, height at withers and previous body condition score record

Chapter submitted for publication as: **Semakula J, Corner- Thomas R.A, R, Morris S.T, Blair, H.T, Kenyon, P.R**. **xxxx.** Predicting ewe body condition score using adjusted liveweight for conceptus and fleece weight, height at withers and previous body condition score record**.** *Translational Animal Science*.

Abstract

The relationship between ewe body condition score (BCS) and live weight (LW) has been exploited previously to predict the former from LW, LW-change and previous BCS records. It was hypothesized that if fleece weight- and conceptus- free live weight and LW-change, and in addition, height at withers were used, the accuracy of current strategies for predicting BCS would be enhanced. Ewes born in 2017 (n = 395) were followed from 8 months to approximately 42 months of age in New Zealand. Individual ewe data were collected on -LW, BCS score at different stages of the annual production cycle (i.e. pre-breeding, at pregnancy diagnosis, pre-lambing and weaning). Additionally, individual lambing dates, ewe fleece weight and height at withers data were collected. Linear regression models were fitted to predict current BCS at each ewe age and stage of the annual production cycle using two LW-based models namely, unadjusted for conceptus weight and fleece weight (LW alone1) and adjusted (LW alone2) models. Further, another two models based on a combination of LW, LW-change, previous BCS and height at withers (combined models) namely, unadjusted (combined1) and adjusted for fleece and conceptus weight (combined2) were fitted evaluated. Combined models gave more accurate (with lower Root Mean Square Error: RMSE) BCS predictions than models based on LW-alone only. However, applying adjusted models did not improve BCS prediction accuracy (or reduce RMSE) or improve model goodness of fit (R^2) ($p > 0.05$). Further, in both LW-alone and combined models, a great proportion of variability in BCS could not be accounted for ($0.25 \geq R^2 \geq 0.83$) and there was substantial prediction error (0.33 BCS ≥ RMSE ≥ 0.49 BCS) across age groups and stages of the annual production cycle and over time (years). Therefore, using additional ewe data which allowed for the correction of LW for fleece and conceptus weight and using height at withers as an additional predictor did not improve model accuracy. In fact, the findings suggest that adjusting LW data for conceptus and fleece weight offers no additional value to the BCS prediction models based on LW. Therefore, additional research to identify alternative methodologies to account for individual animal variability is still needed.

9.1 Introduction

Body condition score (BCS) in sheep is a commonly used subjective measure (Morris et al., 2002; Vieira et al., 2015) to help make flock nutritional and management decisions. Devised by Jefferies (1961) and then revised by Russel et al. (1969), it subjectively quantifies the amount of soft tissue along the lumbar spine (Jefferies, 1961; Kenyon et al., 2014). Body condition score in sheep utilizes a 0.0−5.0 scale range with either half (0.5) units or quarter (0.25) units, and is conducted through the palpation of the lumbar vertebrae immediately caudal to the last rib and above the kidneys (Kenyon et al., 2014).

Body condition score circumvents factors that can confound live weight (LW) such as gut-fill, physiological status, fleece weight and frame size (Coates and Penning, 2000b; Kenyon et al., 2014). Despite the advantages of using BCS over LW to better manage flock nutrition, producers, especially under extensive flock management systems such as in the southern hemisphere, rarely utilise it (Jones et al., 2011; Corner-Thomas et al., 2016). Instead, farmers either depend on inaccurate visual inspection methods or utilise live weight measures only (Besier and Hopkins, 1989). This low uptake among producers is driven by the procedure being subjective; relatively labour intensive and requiring training (Kenyon et al., 2014). Strategies to increase the adoption and use of BCS among producers, such as promotion of producer training and regular assessor recalibration workshops, have not yielded the desired change (Kenyon et al., 2014). This is likely because they do not address how to lessen the additional labour burden related to hands-on BCS, especially in large flocks under extensive management systems. Therefore, it could be reasoned that reliable and precise alternative methods to estimate BCS of sheep that involve reduced hands-on measurement would likely be useful and improve uptake and acceptance of the BCS technique. This indirect method would preferably, be based on already existing and utilized on-farm management tools in order to reduce workload and be easily undertaken and not be subjective in nature.

The relationship between BCS and LW is well established in sheep (Sezenler et al., 2011; Kenyon et al., 2014; McHugh et al., 2019). In Chapter 6, it was demonstrated that BCS is positively and linearly related with LW. This relationship is known to differ by stage of the annual production cycle, age, and breed of ewe (Sezenler et al., 2011; McHugh et al., 2019). This relationship between BCS and LW was utilised to predict current BCS on a 5-point scale from lifetime live weight (current and previous), liveweight change and previous BCS based on linear regression models (Chapter 7). It was demonstrated that with a set of established equations it may be possible to calculate a predicted BCS instantly, at each live weighing, for each sheep. However, a great proportion of variability in BCS remained unaccounted for, leading to less robust models. Further, Chapter 8, machine learning classification algorithms were successfully (with up to 90% accuracy) used to predict BCS using a LW predictors. However, these machine learning classification models were limited to a 3-point scale due to gross class imbalance in BCS data. Full scale BCS (5-point scale: 1.0−5.0) prediction based on linear regression, does not require balanced data. In Chapter 7, it was hypothesized that greater accuracy could be achieved if key variables affecting the relationship between BCS and LW were also accounted for. Morphometric measurements such as height at withers are positively correlated with LW and BCS in sheep (Burke et al., 2004; Holman et al., 2012). Further, pregnancy and fleece weight confound the relationship between BCS and LW (Kenyon et al., 2014; Brown et al., 2015). If these variables could be accounted for, BCS prediction accuracy may potentially be improved. Therefore, the aim of this study was to firstly determine if the ewe BCS prediction accuracies reported in Chapter 7 can be reproduced on in independent dataset and secondly to investigate if the accuracy and scope of BCS prediction equations could be improved by adding information on the height at withers, fleece weight and physiological state of a ewe.

9.2 Material and methods

9.2.1 Experimental design

The current study utilized data collected between 2017 and 2020 from one flock. Romney type ewes were initially raised at Riverside farm (2017−2018) and later (2019) transferred to Keeble farm as part of normal routine farm management. Riverside farm is located 11 km north to north-west of Masterton (40°50'S, 175°37'E) while Keeble farm was 5 km south of Palmerston North (40°24' S and 175°36' E), New Zealand. Ewes were maintained under commercial farming conditions from weaning to 42 months of age (Pettigrew et al., 2018; Pettigrew et al., 2019). A chronological outline of the three-year annual ewe production cycle observed under the current study is summarised in Figure 9.1. A total of 429 ewe lambs born in the same season (Aug-Sep 2017) were followed until maturity at 42 months of age. Data were collected on whether study ewe lambs were born to mature or ewe lambs and in which breeding cycle. Unfasted liveweights and BCS of ewes (born to ewe lambs or mature ewes) were recorded at 6 months of age, pre-breeding (PB), at pregnancy diagnosis (PD), and eight days prior to the start of lambing (PL: pre-lambing) and at weaning (W: Weaning; lambs on average of 3 months of age) in each year. All weight measurement occasions were conducted when ewes were not wet. All ewes were followed for three productive full years. The ewes in this study were themselves presented for breeding at 8 months of age. This study was approved by the Massey University animal ethics committee (protocol number: MUAEC 17/16).

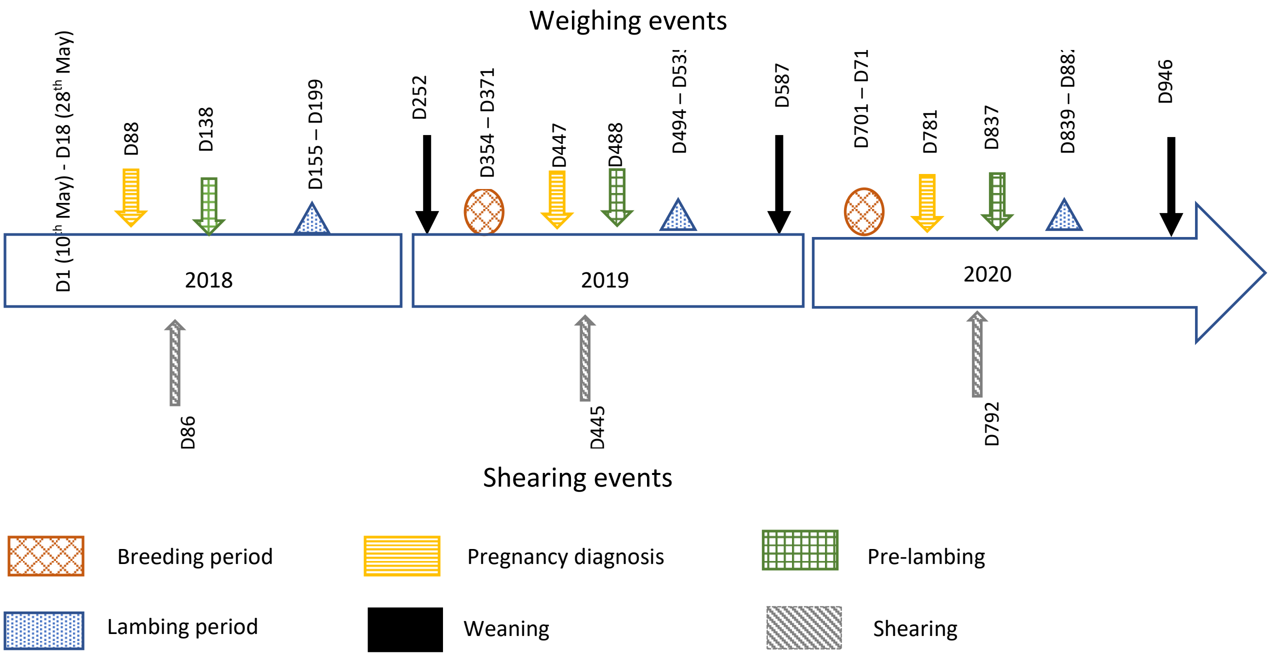


Figure 9.1 Timeline showing the weighing and shearing events during the three-year study. D indicates day of study from 10th May 2018.

All ewes were weighed (to the nearest 0.1 kg) using static digital weighing scales (Tru-Test group, model XR5000). Body condition score was undertaken by one experienced assessor using a 1.0–5.0 scale (1 = thin, 5 = obese) with sheep assessed to the nearest 0.5 of a BCS (Jefferies, 1961; Kenyon et al., 2014). Ewes were shorn each year during late pregnancy (47 to 49 days prior to the start of lambing), and fleece weights were recorded. Estimated fleece weights at the time of the weighing (equation 9.1) in each year were computed by multiplying the annual fleece weight at late pregnancy with the relative proportion of the fleece length (mm) at the corresponding time assuming a shorn fleece length of 150 mm and an amplitude of 19% of the mean (Cottle and Pacheco, 2017).

$$Y_t \text{ (kg)} = Fwt * Rl \qquad 9.1$$

Where Y_t is the estimated fleece weight (kg) at a given time (month), Fwt was the actual fleece weight at the annual shearing (kg), Rl is the proportion of wool length at a given time of the year relative to the wool length when shearing was last done (Length at shearing, mm). The minimum wool length left during shearing was 5.0 mm. All parameters were adapted from Cottle and Pacheco (2017).

The conceptus mass can confound accurate measurement of ewe conceptus free live weight especially from mid-pregnancy onwards (Kenyon et al., 2008; Kenyon et al., 2011b). Adjusted ewe live weight can be obtained if the conceptus mass can be corrected for. Therefore, to allow for the computation of adjusted live weights, lambing dates for each ewe were recorded. The dates were used to estimate days of pregnancy when the live weight measurements were recorded at pregnancy diagnosis (PD) and pre-lambing (PL). The gestation time (days of pregnancy at PD or pre-lambing) was computed as the difference between 147 days (gestation was assumed to be 147 days) and the time from the event (PD or pre-lambing live weight measurement) to lambing. The predicted conceptus and gravid uterus weight was determined using Gompertz equation (equation 9.2) below adapted by Freer et al. (2007). To cater for both single- and twin-bearing ewes, a pooled lamb birth weight (overall weight of both lambs) was computed for twin-bearing ewes.

$$Y = SBW\ exp(A - B(\exp(-Ct)) \qquad 9.2$$

Where Y is the weight of the content of the gravid uterus, SBW is the scaled birth weight (the ratio of the actual birth weight to the standard birth weight of 5 kg assumed by Gompertz equation), t is the gestation length (days) and parameters A, B and C are constants 5.17, 8.38 and 6.08 x 10^{-3}, respectively. A 5 kg lamb at 147 days was used as the standard for scaling of birth weights. The final adjusted ewe live weights excluded fleece weight and gravid uterus weight. Live weights at pre-breeding and weaning were adjusted for fleece weight only, while at pregnancy diagnosis and pre-lambing, both conceptus and fleece weights were adjustedHeight at wither (HW) was recorded every six months using an automatic laser distance measurer (Stanley TLM130i distance meter, max range = 30 m, ± 3mm accuracy) attached to a sliding bar from above the weigh crate (Figures 9.2, 9.3). The height of the ewe was computed using the formula:

$$\text{Unadjusted Height at withers (m)} = X - Z \qquad 9.3$$

Where, X was the distance from the laser meter (X) to the floor of the weigh crate, Z was the distance from the laser meter to the ewe withers. Height at withers was later corrected based on predicted annual fleece growth to generate adjusted HW (Cottle and Pacheco, 2017).

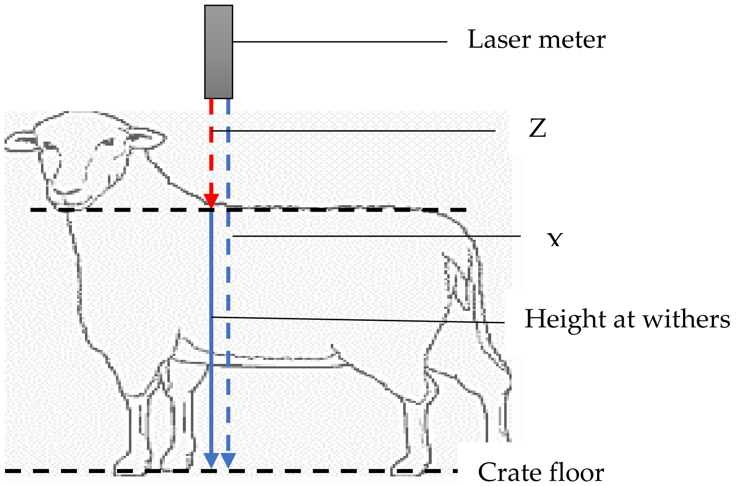


Figure 9.2 Measurement of ewe height at withers

Figure 9.3 laser meter

9.2.2 Statistical analyses

Data were analysed using R program version 3.3.4 (R Core Team, 2016) with package extensions in the caret package (Kuhn, 2008). Similar analytical procedures including variable formulation and selection, model building, cross-validation and evaluation used in Chapter 7, were followed. Consequently, both classification and multiple linear regression approaches were tested. Any missing values were imputed using the preProcess function and bagimput method from the caret package in R (Kuhn, 2008). Additionally, non-numerical data were made numerical and z-transformed (scaled and centred) during analysis using the same preProcess function above. Z-transformed values outside the 95% CI (z ± 1.96 range) were not used in the final analysis. Differences among correlation coefficients were tested for significance based on Fisher's r-to-z transformation. In the present analysis there was high-class BCS imbalance (Table 9.2) making the use of classification methods to predict individual BCS inappropriate (Triguero et al., 2015). In order to predict individual ewe BCS on a full scale (1.0−5.0), an alternative statistically robust method (Norman, 2010) to class imbalance was warranted. Consequently, the multivariate linear model which has been successfully utilised to predict BCS in cattle (Martins et al., 2020) and sheep (Chapter 7) was applied.

9.2.3 Variable selection, model building and validation

The predictors for each BCS were selected through a variable selection technique executed in the R program (R Core Team, 2016) using the elastic net method in the glmnet extension (Friedman et al., 2010) in the caret package (Kuhn, 2008). Models were fitted and validated using a four-step procedure (data partitioning, resampling, model training and validation) as described in Chapters 6 and 7. Using selected predictors regression equations were fitted on a training dataset to predict BCS from lifetime ewe live weight records (current and previous weights), liveweight change (difference in weight between two consecutive weights taken at different time points), height at withers, previous BCS scores (a record of all previous BCS scores) , their lamb birth and weaning weight data in one regression. Initially a total of eleven (11) regression equations (each representing ewe age group and stage of the annual production cycle) were created for BCS prediction based on unadjusted lifetime LW measurements (Liveweight alone1 models). Lifetime measurements were defined as a conglomeration of those ewe measurements taken at both the same and previous time points. A previous measurement was that taken at a different time point (different stage of the annual production cycle) prior to the current one. Liveweight change refers to the change in live weight between two time points. Further, 11 more equations were generated incorporating liveweight change and previous BCS in addition to lifetime live weight (combined1 models). The process of generating BCS prediction equations above was repeated based on adjusted LW (adjusting for conceptus weight and fleece weight) (Liveweight alone2 models) and based on adjusted LW, liveweight change, height (adjusted for fleece growth) at withers and previous BCS (combined2 models). A description of variables is given in Table 9.1.

Table 9.1 Explanation of live weight (LW), liveweight change, height at withers (H) and body condition score (BCS) variables by ewe age group and stage of the annual production cycle.

Age (Months)	Stage of the Annual production cycle	LW*	BCS§	Change in Live weight£	HW€
8–18	Pre-breeding	WP1	BP1		PH1
	Pregnancy diagnosis	WD1	BD1	WT11(WD1–WP1)	DH1
	Pre-lambing	WL1	BL1	WT12(WL1–WD1)	LH1
	Weaning	WW1	BW1	WT13(WW1–WL1)	WH1
19–30	Pre-breeding	WP2	BP2	DW-T1(WP2–WW1)	PH2
	Pregnancy diagnosis	WD2	BP2	WT21(WD2–WP2)	DH2
	Pre-lambing	WL2	BL2	WT22(WL2–WD2)	LH2
	Weaning	WW2	BW2	WT23(WW2–WL2)	WH2
31–42	Pre-breeding	WP3	BP3	DW-T2(WP3–WW2)	PH3
	Pregnancy diagnosis	WD3	BP3	WT31(WD3–WP3)	DH3
	Pre-lambing	WL3	BL3	WT32(WL3–WD3)	LH3
	Weaning	WW3	BW3	WT33(WW3–WL3)	WH3

LW*; live weight at pre-breeding (WP), pregnancy diagnosis (WD), pre-lambing (WL), and weaning (WW). BCS§; body condition score at pre-breeding (BP), pregnancy diagnosis (BD), pre-lambing (BL), and weaning (BW). Change in live weight£: WT; change in live weight between successive measurements within age groups, DW-T; change in live weight between successive measurements between age groups. HW€; Height at withers at pre-breeding (PH), pregnancy diagnosis (DH), pre-lambing (LH), and weaning (WH).

9.2.4 Model evaluation

Models were evaluated as described in Chapter 7. Model performance evaluation was conducted on training dataset using two metrics (Theil, 1958; Botchkarev, 2019) adjusted coefficient of determination (adj. R^2) and the root mean square error (RMSE). Each BCS prediction model validation was conducted on the testing dataset, with each replicated 1000-fold. The quality and success of the prediction models was assessed using the coefficient of determination (r^2), mean bias, root mean squared error (RMSE), residual prediction deviation (RPD), the ratio of performance to interquartile distance (RPIQ) and percent error (RPE) (McDowell et al., 2012), overall adjusted R^2 value and error metrics between models, were compared based on Wilcoxon signed-ranks test (Conover, 1973; Rahe, 1974) and a two-tailed paired t-test (Kim, 2015).

9.3 Results

9.3.1 Descriptive statistics

The frequency of ewe BCS score across age group and stage of the annual production cycle is presented in Appendix X Table 1. The majority of the ewes had BCS ranging from 2.5 to 3.0, while the extreme BCS scale values (1.5 and 5.0) were the least common. Within age groups, the most frequent ewe BCS at 8−18 months was 2.5 across stages of the annual production cycle, at 19−30 months was 3.0 across all stages of the annual production cycle except at weaning and at 31−42 months there was no clear pattern.

Summaries of ewe LW, BCS and HW from 8 to 42 months of age are presented in Table 9.2. Both BCS and HW did not significantly change ($p > 0.05$) over time and across stages of the annual production cycle, LW varied ($p < 0.05$) with annual production cycle and increased with ewe age. Unadjusted LW continued to increase with ewe age beyond 30 months. However, adjusted live weight increased with age up to 30 months before plateauing.

Table 9.2 Mean live weight unadjusted and adjusted for conceptus and fleece weight (LW), height at withers (HW) and body condition score (BCS) with respective standard deviations by ewe age group and stage of annual production cycle.

			LW		HW		BCS
Age (months)	Stage of annual production cycle	n	Unadjusted	Adjusted	Unadjusted	Adjusted	
8–18	Pre-breeding	428	43.7 (5.61)	41.5 (5.46)			2.8 (0.42)
	PD	429	48.8 (5.83)	45.7 (5.42)			2.7 (0.39)
	Pre-lambing	428	52.6 (7.49)	52.0 (7.47)	0.61 (0.032)	0.58 (0.032)	2.8 (0.41)
	Weaning	429	59.7 (7.10)	58.6 (7.05)			2.8 (0.53)
19–30	Pre-breeding	427	62.8 (6.67)	59.1 (6.73)	0.61 (0.038)	0.59 (0.038)	3.0 (0.61)
	PD	426	63.0 (7.09)	60.2 (6.74)	0.60 (0.036)	0.58 (0.036)	3.3 (0.63)
	Pre-lambing	424	70.8 (7.70)	62.0 (6.60)			3.2 (0.63)
	Weaning	424	66.1 (8.67)	64.2 (8.67)	0.63 (0.033)	0.59 (0.033)	2.8 (0.67)
31–42	Pre-breeding	401	68.9 (7.71)	66.4 (7.74)			
	PD	402	70.7 (7.76)	64.8 (7.57)	0.62 (0.047)	0.59 (0.033)	3.1 (0.63)
	Pre-lambing	399	88.8 (9.32)	64.3 (8.27)			3.4 (0.65)
	Weaning	402	69.0 (9.74)	66.8 (9.70)	0.64 (0.033)	0.61 (0.047)	2.8 (0.78)

Values in parenthesis indicate the standard deviation. Adjusted indicates that variables were corrected for fleece conceptus weight (LW) and fleece growth (LW and HW).

9.3.2 Correlation between live weights and height at withers

The relationship between ewe live weight (LW) and height at withers (HW) was positive but weak to moderate across age groups and stages of the annual production cycle, regardless of whether unadjusted or adjusted LW was used (Appendix X Table 2). However, a negative association between LW and HW was observed at 19−30 months at pre-breeding. There was no pattern in the strength of BCS-HW association between same and different time points.

9.3.3 Correlation between BCS and Live weights

There was a linear association between LW and BCS in all age groups and stages of the annual production cycle, but the association was weak to moderate, regardless of whether unadjusted or adjusted LW was used (Appendix X Table 3). Further, this association was comparable ($p > 0.05$) for both unadjusted and adjusted LW. Both the weakest and strongest relationships were observed at weaning. The relationships, however, were strongest when live weight and BCS measurements were from the same time point (pair of LW-BCS measurements taken at the same time) except at pre-lambing 8−18 months, compared with when lifetime (i.e., measurements taken at different time points) records were used..

9.3.4 Correlation between BCS and height at withers

Generally, there was a poor linear association between ewe HW and BCS in all age groups and stages of the annual production cycle, regardless of whether unadjusted or adjusted HW (Appendix X Table 4). At any one time point, the relationship between BCS and HW did not vary ($p > 0.05$) across age and stage of the annual production cycle except for 19−30 month ewes at weaning ($p < 0.01$) and 31−42 month ewes at pre-lambing ($p < 0.01$) and weaning ($p < 0.05$). There was no clear pattern in the change of strength of BCS-HW association over time.

9.3.5 Coefficient of Determination (R^2) and Number of Predictors

To predict BCS at any given time, all current and previous individual live weights (liveweight alone models) were included in linear regression equations. Separate models were formulated for unadjusted and adjusted LW (based on training dataset). The adjusted R^2 values averaged across folds 0.38 (0.10 to 0.74), regardless of the time point. The adjusted R^2 values were comparable ($z = 0.37$, $t_{10} = 0.56$, $p > 0.05$) for both adjusted and unadjusted BCS prediction models across age groups and stages of the annual production cycle (Figure 9.4). However, the average adjusted R^2 value was greater for unadjusted than adjusted LW models ($z = 2.40$, $t_{10} = 2.23$, $p < 0.05$). Within age groups, across stages of the annual production cycle, adjusted R^2 value varied with no clear pattern (Figure 9.4). There was a trend for adjusted R^2 to improve at older ages, when a greater amount of previous live weight information was known. In general, the adjusted R^2 value was highest at weaning with no clear pattern in the lowest value. The average number of live weight predictors (significant variables) for BCS prediction was comparable for models using unadjusted as well as adjusted LW (Average: 6, 1 to 11), and increased with ewe age (Figure 9.4).

To improve the prediction of current BCS, the LW alone models were expanded by adding the unadjusted LW difference (change in live weight measurements from adjacent time points) and all preceding BCS (combined unadjusted models) or by adding the adjusted LW differences and height at wither, and all preceding BCS (combined adjusted models). The overall proportion of variance explained (adjusted R^2) improved ($z=3.62$, $t_{21} = 5.71$, $p < 0.001$) by approximately 1.3 times (from 0.38 to 0.50) in all combined model categories compared with LW models (Figure 9.4). However, the adjusted R^2 values were comparable ($z = 1.07$, $t_{10} = 0.99$, $p > 0.05$) for both adjusted and unadjusted models across age groups and stages of the annual production cycle (Figure 9.4). Further, the adjusted R^2 values were marginally greater in combined models than liveweight alone models across age and stages of the annual production cycle. The highest adjusted R^2 values were achieved at the weaning period with no clear pattern concerning the lowest value. The number of significant predictors for BCS was higher (average: 10, from 1 to 16 for unadjusted and 1 to 21 for unadjusted) in the combined models compared with liveweight alone models (Figure 9.5). Overall, the number of predictors was increased 1.5 and 2.0 times for unadjusted and adjusted combined models, respectively, compared with LW alone models.

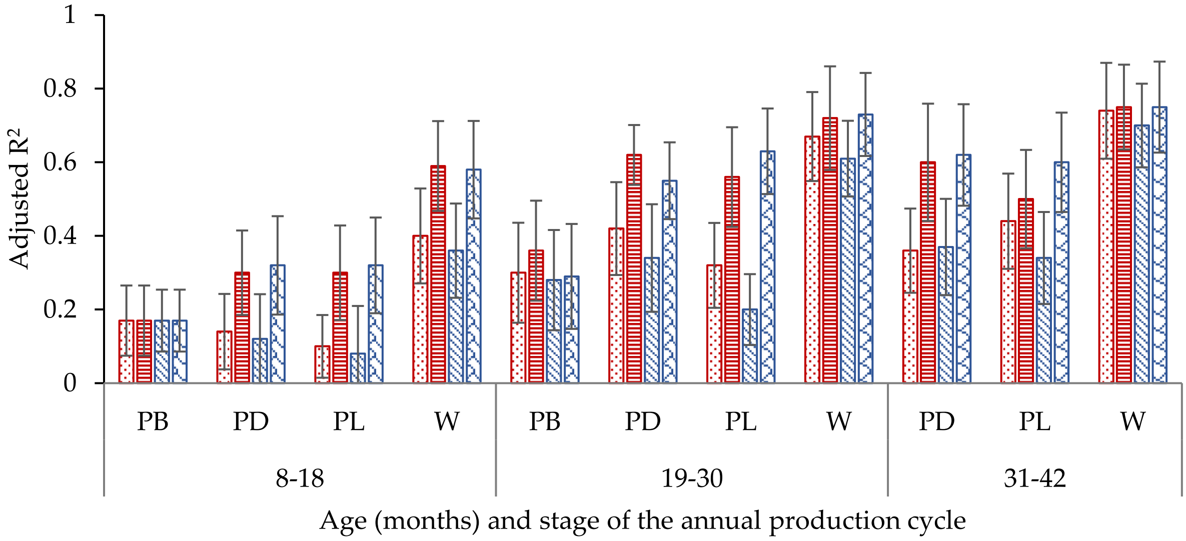


Figure 9.4 Adjusted coefficient of variation (Adjusted R^2, with standard deviations) of models (dotted bar: unadjusted liveweight alone models, horizontal stripes: combined models based on unadjusted LW, liveweight change and previous BCS, diagonal stripes: adjusted liveweight alone, shingled: adjusted live weight, liveweight change, height at withers and previous BCS) for current BCS prediction across the stage of the annual production cycle and ewe age group. Colours (Red indicates unadjusted live weight while blue indicates adjusted liveweight was used). PB, PD, PL, W indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively. In large samples where bootstrapping is applied, the standard deviations approximate the standard errors.

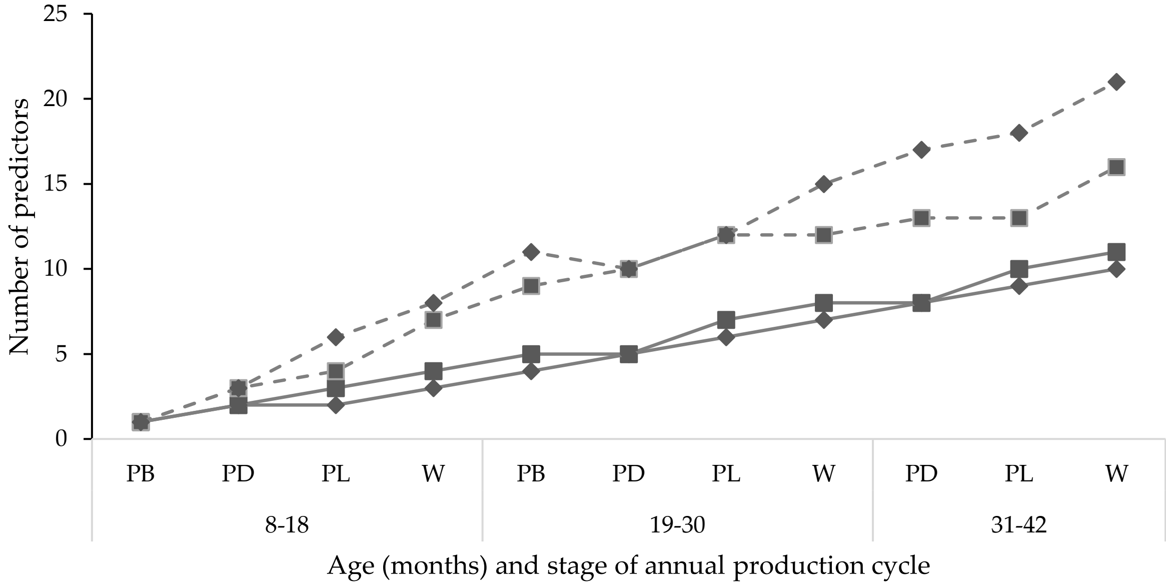


Figure 9.5 Number of predictors of models (solid lines with squares: unadjusted liveweight alone models, solid line with diamonds: adjusted liveweight alone models, dashed line with squares: combined models based on unadjusted LW, liveweight change and previous BCS, dashed line with diamonds: adjusted live weight, liveweight change, height at withers and previous BCS for BCS prediction at given time across the stage of the annual production cycle and ewe age group. PB, PD, PL, W indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively.

9.3.6 Prediction accuracy

To access the accuracy of predicting BCS, several prediction error metrics (MAE, RMSE, RPE) were computed. The error metrics appeared to vary across ($p < 0.05$) age group but not ($p > 0.05$) stage of the annual production cycle except for 19−30-month-old ewes, when live weight or combined models were used to predict BCS (Tables 9.3, 9.4, Appendix X Figure 1). Using adjusted LW did not affect BCS prediction accuracy (±2SD, $p > 0.05$) except for the 19−30-month-old ewes at pre-lambing. The average prediction error associated with BCS prediction in terms of MAE and RMSE were 0.38 and 0.45, and 0.32 and 0.40 body condition scores for liveweight alone and the combined models, respectively. In adjusted models, the average prediction error associated with BCS prediction in terms of MAE and RMSE were 0.37 and 0.45, and 0.33 and 0.41 body condition scores for liveweight alone and the combined models, respectively. However, combined models improved ($z = 5.41$, $t_{21} = 2.08$, $p > 0.001$) the BCS prediction error by 10.7% (Average RMSE: 0.45 vs 0.40) compared with LW alone models.

The magnitude of the BCS prediction error was moderate to high in both the live weight and combined models, based on the smallest unit of measurement (0.5). The BCS predictions using the unadjusted liveweight alone and combined models were, on average, 15.4% and 13.5%, respectively, from the actual value. In adjusted models, the predictions deviated by 15.9% and 13.4% respectively, for LW alone and combined models. Therefore, combined models improved the BCS prediction error prevalence by 9.6% compared with LW alone models.

Models were categorized regardless of model type as weak (RPD: 1.06 to 1.35) or very poor to fair (RPIQ: 1.47 to 1.85). There was inconsistency in the BCS prediction model performance where a model with relatively good RPD (>1.4) had a poor RPIQ (<1.4) and vice versa. Using adjusted LW or unadjusted LW did not affect ($p > 0.05$) both model RPD and RPIQ metrics. However, both RPD and RPIQ were improved ($p < 0.05$) by 10 to 16% in the combined than LW alone models.

Table 9.3 Coefficient of determination (r^2), bias, root mean square error (RMSE), mean absolute error (MAE), relative prediction error (RPE) residual prediction deviation (RPD), and the ratio of performance to interquartile distance (RPIQ) based on testing data for the prediction of BCS in ewes between 8 and 42 months by stage of the annual production cycle using unadjusted live weight and adjusted live weight (LW) alone models.

	8–18				19–30				31–42		
Metric	PB	PD	PL	W	PB	PD	PL	W	PD	PL	W
(a) Liveweight alone[1] models (Unadjusted)											
r^2	12.90	13.89	10.30	36.70	25.50	26.61	17.50	64.02	33.20	20.33	71.10
BIAS	0.007	−0.043	0.004	−0.013	−0.05	−0.015	0.02	−0.02	0.204	−0.047	−0.152
RMSE	0.39	0.37	0.39	0.43	0.53	0.55	0.54	0.38	0.50	0.49	0.44
MAE	0.32	0.3	0.32	0.33	0.43	0.45	0.46	0.31	0.43	0.44	0.35
RPE	14.90	15.01	13.21	16.20	16.03	15.30	14.70	13.20	16.00	14.30	15.80
RPD	1.14	1.06	1.07	1.26	1.32	1.27	1.30	1.71	1.26	1.26	1.83
RPIQ	1.25	1.25	1.39	1.16	1.04	1.04	1.02	1.32	1.00	2.04	1.14
(b) Liveweight alone[2] models (Adjusted)											
r^2	12.30	15.81	13.50	36.78	32.67	26.70	32.40	68.31	44.16	34.60	57.60
BIAS	0.006	0.019	0.003	−0.088	−0.002	−0.006	−0.008	−0.003	−0.037	0.038	0.063
RMSE	0.4	0.37	0.38	0.41	0.49	0.54	0.54	0.41	0.48	0.52	0.48
MAE	0.32	0.30	0.31	0.33	0.43	0.44	0.43	0.32	0.41	0.43	0.37
RPE	10.6	12.95	13.38	14.75	16.33	16.12	16.72	14.64	15.53	15.25	17.2
RPD	1.08	1.09	1.07	1.24	1.23	1.17	1.19	1.79	1.35	1.25	1.53
RPIQ	1.25	1.35	1.32	1.22	2.04	1.39	0.93	1.22	2.08	1.92	1.04

PB, PD, PL, W indicate the four stages of the annual production cycle including pre-breeding, pregnancy diagnosis, pre-lambing, and weaning, respectively. Interpretation of measures: The best model has the highest r^2, RPD, and RPIQ, and the lowest RMSE and RPE. Ranges for values: r^2 (0: Indicates that the model accounts for none of the variability of the response data around its mean, 1.0 indicates that the model accounts for all the variability). RPD (< 1.4: Weak, 1.4 < RPD < 2.0: Reasonable, > 2.0: Excellent). RPIQ (< 1.4: Very poor, 1.4 < RPIQ < 1.7: Fair, 1.7 < RPIQ < 2.0: Good, 2.0 < RPIQ < 2.5: Very good, > 2.5: Excellent). (a) (b) superscripts [1,2] indicate model based on unadjusted or adjusted live weight, respectively. Bias (Positive value indicates overestimation; negative sign indicates underestimation). Adjusted indicates that a model was based on live weight corrected for conceptus and fleece weight.

Table 9.4 Coefficient of determination (r^2), bias, root mean square error (RMSE), mean absolute error (MAE), relative prediction error (RPE) residual prediction deviation (RPD), and the ratio of performance to interquartile distance (RPIQ) based on testing data for the prediction of BCS in ewes between 8 and 42 months by stage of the annual production cycle using unadjusted and adjusted combined models.

	8–18				19–30				31–42		
Metric	PB	PD	PL	W	PB	PD	PL	W	PD	PL	W
(a) Combined[1] models (Unadjusted)											
r^2	12.9	31.3	26.6	51.5	29.9	55.1	58.3	68.8	54.7	54	71
BIAS	0.007	−0.025	−0.005	0.009	0.051	−0.038	0.007	0.065	−0.014	−0.155	0.011
RMSE	0.39	0.33	0.34	0.36	0.49	0.47	0.40	0.39	0.42	0.45	0.41
MAE	0.32	0.27	0.28	0.29	0.41	0.31	0.31	0.31	0.34	0.35	0.31
RPE	14.90	12.13	11.97	12.86	16.69	12.18	12.42	14.08	13.64	13.20	14.96
RPD	1.14	1.20	1.18	1.42	1.20	1.48	1.56	1.73	1.50	1.49	1.84
RPIQ	1.25	1.52	1.47	1.39	2.00	1.83	2.50	1.28	1.19	1.11	1.22
(b) Combined[2] models (Adjusted)											
r^2	13.0	31.9	18.7	53.7	36.4	55.5	57.1	67.2	51.7	57.8	71.3
BIAS	0.006	−0.001	−0.002	0.021	−0.01	0.005	0.054	−0.034	0.043	0.031	−0.054
RMSE	0.40	0.33	0.39	0.35	0.49	0.46	0.41	0.42	0.44	0.42	0.40
MAE	0.33	0.28	0.27	0.30	0.43	0.35	0.32	0.32	0.34	0.34	0.31
RPE	10.6	11.76	13.64	12.68	16.33	13.65	12.58	15.16	14.15	12.35	14.55
RPD	1.08	1.24	1.08	1.47	1.25	1.34	1.54	1.79	1.40	1.55	1.85
RPIQ	1.25	1.56	1.28	1.43	2.04	1.63	1.22	1.19	2.27	2.38	1.25

PB, PD, PL, W indicate the four stages of the annual production cycle including pre-breeding, pregnancy diagnosis, pre-lambing, and weaning, respectively. Interpretation of measures: The best model has the highest r^2, RPD, and RPIQ, and the lowest RMSE and RPE. Ranges for values: r^2 (0: Indicates that the model accounts for none of the variability of the response data around its mean, 1.0 indicates that the model accounts for all the variability). RPD (< 1.4: Weak, 1.4 < RPD < 2.0: Reasonable, > 2.0: Excellent). RPIQ (< 1.4: Very poor, 1.4 < RPIQ < 1.7: Fair, 1.7 < RPIQ < 2.0: Good, 2.0 < RPIQ < 2.5: Very good, > 2.5: Excellent). (a) Unadjusted indicates that models were based on all previous and current crude and previous live weights, liveweight changes and previous BCS). Adjusted indicates that models were based on all previous and current live weights and liveweight changes corrected for conceptus and fleece weight, adjusted height at withers, and previous BCS. The superscripts [1,2] indicate without and with adjusted HW in the model, respectively. Bias (Positive value indicates overestimation; negative sign indicates underestimation).

9.4 Discussion

The aim of this study was to explore the possibility of improving the prediction accuracy of BCS using a ewe's production characteristics as they aged from eight through to approximately forty-two months. This was a follow-up study to Chapter 7. Previously, using a different dataset, the relationship between live weight and BCS at a given time point, and the possibility of using a linear combination of a ewe's unadjusted lifetime LW, liveweight change and previous BCS data to predict BCS at a given time, were examined (Chapter 6 and 7). Weak to moderate levels of BCS prediction accuracy were achieved. It was then postulated that if corrected live weights (corrected for conceptus and fleece weight) and wither height (corrected for fleece length) data were used, BCS prediction accuracy would be improved.

In this study the majority of the ewes had BCS between 2.5 and 3.0 which falls within the recommended BCS range (2.5–3.5) for optimal productivity (Kenyon et al., 2014). Additionally, there were few thin or obese ewes in the 8 to 18-month-old group. These observations combined indicate that ewes were supplied with sufficient nutritional requirements through their first reproductive cycle. Further, this study demonstrated unadjusted LW continued to increase beyond 30 months of age but adjusted LW (adjusted for conceptus and fleece weight) plateaued. The observed trend in adjusted LW corroborates an earlier study which reported that mature Romney ewe weight was achieved by 33 months (Pettigrew et al., 2019). It appears that the confounding effects of conceptus and fleece weight increase with age, causing the apparent increase in weight unadjusted LW.

This study showed a linear relationship between LW and HW. This relationship was generally positive for most stages of the annual production cycle and age groups. It was, however, not clear why this relationship was negative for 19−30-month-old ewes at pre-lambing. Prior to breeding, farmers enhance their feeding strategies in a process known as flushing to ensure as many ewes reach the required breeding weight regardless of their frame sizes (Kenyon et al., 2011b). Given that fact that this was the same cohort of ewes, it is possible that changes in nutritional effects could have randomly altered the relationship between LW and HW. With the moderate strength of association between LW and HW, height at withers, was expected to significantly affect the relationship between LW and BCS. However, HW was poorly correlated with BCS. There was a weak to moderate correlation between LW and BCS as reported in Chapter 6.

The observation that LW alone models were not as good as combined ones and, thus, likely to be unreliable in predicting future BCS based on linear regression, corroborates our previous findings (Chapter 7). The variability in BCS explained for both live weight and combined models increased with the number of predictors in the model. This was expected, as it is known that as the number of predictors that significantly relate to the dependent variable increase, the proportion of the variance due to the regression increases (Li, 2017). However, in this study, a considerable amount of variability in BCS ($0.26 \leq R^2 \leq 0.83$ and $0.25 \leq R^2 \leq 0.72$) remained unaccounted for in both liveweight alone models and combined models, respectively. Some of the reasons for the apparent failure for both liveweight alone and combined models to account for more of the variability in BCS include, (i) assessor consistency over time, (i) losses in live weight due to gut-fill and urination when ewes were weighed at different times, (iii) confounding effects of fleece weight, and conceptus weight (Chapter 7). The consistency between BCS assessors varies between from 5% to 27% and 40% to 60%, and within assessors from 16% to 44% and 60% to 90% for inexperienced and experienced assessors, respectively (Kenyon et al., 2014). The current study a single experienced assessor (with more than 30 years of experience in BCS assessment) was used to determine all BCS to ensure consistency. It is, therefore, unlikely that the data used in this study was affected by assessor reliability. Liveweight losses resulting from fluctuations in gut-fill can account for between 5% and 23% of total live weight in ruminants (Hughes, 1976; Moyo and Nsahlai, 2018). Thus, the duration between feeding and recording an individual's live weight can influence the accuracy of the live weight. Further, ewe fleece weight, pregnancy and lambing data were collected and used to correct LW. Given that standard equations, with little known error rates and repeatability were used to adjust live weight, it is possible that these strategies could have introduced some error cancelling the effect of adjusting for LW confounders. The study did not measure individual time off feed prior to weighing, a function that many electronic weighing systems now have the potential to account for. Future studies should examine if the accuracy of BCS prediction can be improved by accounting for gut-fill fluctuations. In regression models all residual error is assumed to be contributed by the predictors and thus, any inaccuracies in their measurement should be of concern (Dosne et al., 2016). Losses in live weight due to gut-fill changes and urination in relation to when ewes were weighed and the effect of pregnancy on live weight are therefore, of concern, as they affect live weight a key variable for BCS prediction. When predictor variables are imprecise, estimations based on the standard model assumptions can lead to inaccurate parameter estimates even when large samples are used (Hausman, 2001; Pischke, 2007). Therefore, if errors in the measurement of live weight could be minimized, the goodness-of-fit and accuracy of BCS prediction models should increase. In delayed weighing, accounting for liveweight losses with respect to time of delay (the duration from when the animal last fed to weight recording) using prediction equations, offers a practical solution. These time-dependent, live weight adjusting equations for ewes have been developed but are not regularly used (Burnham et al., 2009; Wishart et al., 2017).

The BCS prediction models using liveweight alone had larger error (MAE and RMSE) and lower RPD and RPIQ values, compared with combined models which led to high relative error prevalence (RPE). Combined models reduced the magnitude of all the prediction error metrics but were greater than those observed in our previous study (Chapter 7). The model BCS prediction percentage error (RPE) was above the desired 10% (Hagerman et al., 2017; Lalic et al., 2018). The large BCS prediction error values (60 to 108% of the smallest unit on a 0.5 decimal scale) in the present study (where a 0.5-unit change in BCS changes the performance rank of a ewe) could lead to inaccurately predicted BCS values, thereby, greatly influencing management decisions. Ideally, all prediction models should have had resolutions as low as 0.02 body condition scores. However, due to the intractable discrete nature of the BCS scale used, such resolutions cannot be achieved (Chapter 7). It has been suggested that decisions concerning strategic feeding and management of ewes to maximise performance should be based on a critical range of BCS values (i.e., 2.5 to 3.5) (Kenyon et al., 2014). The predictions found in this study may, therefore, overestimate or underestimate measures by 0.33 to 0.54 BCS, which could substantially change the ranking of a ewe, leading to less robust management decisions, which in turn could reduce flock productivity. The greater BCS prediction error than reported in our previous study (Chapter 7) could be explained by the smaller sample size used in the current study leading to greater variability in the outcome and predictor variable measurements.

The findings suggest that using quantitative traits (physical and linear morphometric measurements) may not be sufficient to predict sheep BCS on a full range scale (1.0−5.0). Therefore, further studies us data such as image analysis (Computed Tomography: CT scans, dual-energy X-ray absorptiometry: DXA), and automated metabolic profiles to account for individual animal variability may be warranted. Where a narrow range of BCS such as 1.0−3.0 is acceptable, further research should look at extending machine learning algorithms across all age groups and stages of the annual production cycle. Given the limitations of predicting BCS, itself a predictor of body composition. It would be worthwhile investigating how accurately live weights and other predictors would predict total body fat and muscle weights, or proportions given they are more objective and continuous variables. The first step in these types of studies would require animals to be euthanized and/or tools such as CT scans.

9.5 Conclusion

The combined models improved the proportion of variability in BCS that could be accounted for, as well as the accuracy metrics across all age groups and stages of the annual production cycle and over time (years), compared with the liveweight alone models. Using ewe data to correct LW (correct for fleece weight and conceptus weight) and height at withers as additional predictor did not offer better model accuracy. The most common ways of determining BCS is through a direct hands-on method, however, if it is not possible, the equations generated in the current and previous study (Chapter 7) could be used to predict BCS. These equations could potentially be incorporated in electronic weighing systems that utilize lifetime data especially in large extensively run sheep flocks. However, the 30% to 90% variability in BCS that was unaccounted for, even in the combined models, coupled with the large prediction error associated with our equations dictates that they should be used with caution. Additional ways of accounting for individual variability in BCS could ameliorate the accuracy of BCS and warrant investigation.

Chapter 10. General discussion

10.1 Introduction

Live weight (LW) is a broadly accepted proxy for the energy status of sheep at a given time, while change in live weight is indicative of whether it is in either a positive energy balance (liveweight gain) or a negative energy balance (liveweight loss) (Young and Corbett, 1972; Brown et al., 2005; Wishart et al., 2017). Therefore, live weight provides a basis for decision making regarding sheep management, therefore, accurate determination of LW is important. Live weight measurements can be affected by a number of factors including: growth, nutrition, health, stress, frame size, fleece weight, physiological state and genotype (Kenyon et al., 2014; Brown et al., 2015). Further, the contents of the rumen (fluid and feed) can account for between 5 and 23% of total live weight in ruminants (Hughes, 1976; Moyo and Nsahlai, 2018). Varying levels of weight loss (1.5 to 10% of initial live weight) have been reported within flocks while waiting to be weighed (Hughes, 1976; Burnham et al., 2009; Wilson, 2014; Wishart et al., 2017). These levels of liveweight loss are likely to interfere with a comparison of live weight particularly when small liveweight changes are being investigated or when live weight is used to make decisions are based on thresholds. Existing strategies to reduce liveweight variation have been limited to standardizing the weighing protocol (Coates and Penning, 2000a; Wishart et al., 2017). Such methodologies to reduce variation are cumbersome, time consuming and, therefore, not generally utilised except in experimental situations. Therefore, new approaches to determine and adjust for variations in live weight between animals and specific periods of time when sheep do not have access to feed and water while waiting to be weighed need investigation. The development of these approaches will require an understanding of the factors influencing liveweight loss.

Body condition score (BCS) is an alternative but subjective measure which provides an estimate of an animal's soft tissue reserves, predominantly fat, and is used widely by farmers and researchers to determine the physiological state of an animal

(Morris et al., 2002; Vieira et al., 2015). Body condition score can circumvent the shortcomings of LW, which include the effect of gut-fill, frame size, fleece weight and physiological state (Kenyon et al., 2014; Brown et al., 2015; Morel et al., 2016). Further, body condition score can be easily learned and is cost-effective and requires no specialized equipment (Kenyon et al., 2014). Despite the advantages of using BCS over live weight to better manage flocks, it is uncommon for farmers (7−40%, only) especially in extensive production systems to regularly and objectively do so (Jones et al., 2011; Corner-Thomas et al., 2016). The reasons for low BCS uptake among farmers include the subjective nature, labour burden and constant recalibration of assessors (Kenyon et al., 2014). Strategies to increase the adoption and use of BCS among farmers and the reliability of measures have been limited to promotional workshops and hands-on training (Kenyon et al., 2014). However, these strategies do not directly address how to reduce the labour burden associated with hands-on BCS. Therefore, it is argued that, consistent and accurate alternative methods to estimate body condition score of sheep that require less hands-on measurement would likely be advantageous and improve uptake and use. These would be based on a management tool already utilized on farm, to reduce workload, be quick and not subjective in nature.

The aims of this thesis were to gain a clearer understanding of the factors that influence the rate of liveweight loss of fasting ewes, to derive equations that improve the measurement of live weight measurement, to understand the factors affecting the relationship between live weight and BCS and to develop BCS prediction equations based on a ewe's characteristics.

10.2 Chapter summaries (summary of main findings and conclusions drawn)

In this Chapter (Chapter 10), the general outcome of the experiments, their results, conclusions, and implications for live weight and BCS measurement are discussed. In addition, limitations and weaknesses of the research are identified and discussed. This

Chapter ends with an overview of the main conclusions of the research and recommendations for future research.

10.2.1 *Chapter 3. The effect of herbage type on the rate of liveweight loss of fasting ewe lambs*

In this Chapter, it was hypothesised that liveweight loss rate in fasting ewe lambs would be lower when offered ryegrass-based swards than herb-clover-based swards prior to fasting. It was found that ewe lambs offered the herb-clover mix achieved liveweight loss rates 2.0 times greater than ewe lambs offered ryegrass-clover based swards. Although the study did not perform feed chemical analysis, the higher liveweight loss in ewe lambs offered the herb-clover based swards than ryegrass-clover based swards was attributed to likely higher concentration of readily fermentable soluble sugars and pectin, and lower concentrations of cellulose and hemicellulose than grass-based diets (Barry et al., 1999; Moyo and Nsahlai, 2018). This study confirmed the suggestion that prior to weight measurement, it is important to have previously fed the animals the same ration to eliminate ration effects on rumen gut-fill (Meyer et al., 1960).

10.2.2 *Chapter 4. The effect of herbage availability and season of year on the rate of liveweight loss of fasting ewe lambs*

Previous studies on fasting ewe lamb liveweight change studies utilised one diet type with no indication of herbage quantity offered or season. This Chapter, therefore, investigated the effect of herbage availability and season on the rate of ewe liveweight loss was examined and correction equations for delayed live weights developed for use under commercial conditions. It was found that the rate of liveweight loss increased with herbage availability. Further, this rate of liveweight loss was greater in winter than autumn. The higher liveweight loss rate in ewe lambs offered the High diet and lower rate in lambs on the Low diet was due to the consistently lower percentage dry matter (% DM) in the former and vice versa. The higher liveweight loss rate in winter than autumn in the

Medium and High diets was attributed to the seasonal differences in the chemical composition of the feeds. Applying live weight correction equations on delayed live weight data provided more accurate estimates (33 to 55%) of “without delay” live weight than using the delayed live weight.

Combined these results suggest that beyond grazing ewe lambs on the same diet type and weighing them “without delay”, the quantity of herbage and season should be considered when weighing their ewe lambs. Where “without delay” live weights are not achievable, the correcting equations developed in this Chapter should be used to obtain more accurate “without delay” live weight estimates. These correction equations could be incorporated into weighing systems to automatically give real time adjusted ewe lamb live weights.

10.2.3 Chapter 5. The effect of herbage availability and ewe physiological state, stage of pregnancy and pregnancy-rank on the rate of liveweight loss of fasting mixed-aged ewes

Previously, physiological state has been reported to impact the rate of liveweight loss in ewes offered a fixed narrow range of herbage mass (Burnham et al., 2009). It is possible that by varying the quantity of herbage offered to ewes, their rates of weight loss would also vary. This study, therefore, investigated the effect of herbage availability and physiological state (non-pregnant vs pregnant), stage of pregnancy and pregnancy-rank) on the rate of ewe liveweight loss was examined and correction equations for delayed live weights developed for use under commercial conditions. It was found that the rate of ewe liveweight loss was greater in ewes offered the High than the Low herbage level across physiological state. Further, this rate of liveweight loss was greater in ewes at approximately 100 than 130 days of pregnancy. The observation that the rate of weight loss was greater in ewes offered the High herbage level agrees with the findings in Chapter 4. The lower liveweight loss rate at 130 days of pregnancy has been attributed to the decrease in the reticulo-rumen volume during advanced pregnancy (Forbes, 1969). When correction equations were applied to adjust for ewe liveweight loss, there was increase in accuracy (58 to 67%) in “without delay” live weight estimates.

Combined these results suggest that in addition to weighing ewes “without delay”, grazing them on the same diet type and quantity prior to weighing as measures to ensure accurate live weight measurement, ewe physiological state should also be considered. Where “without delay” live weights are no achievable, the correction equations developed in this Chapter should be used to obtain more accurate “without delay” ewe live weight estimates. These correction equations could be incorporated into weighing systems to automatically give real time adjusted ewe live weights.

10.2.4 Chapter 6. The effect of age, stage of the annual production cycle and pregnancy-rank on the relationship between live weight and body condition score of a ewe

In this Chapter, it was hypothesised that the relationship between LW and BCS in Romney ewes would vary by ewe age, of stage of the annual ewe production cycle and pregnancy-rank. It was found that the relationship between LW and BCS increased with ewe age and differed by stage of the annual ewe production cycle and pregnancy-rank. Further, this relationship between LW and BCS was found to be sufficiently described by the simple linear regression model as reported in many studies (Kenyon et a., 2014; Morel et al., 2016; McHugh et al., 2017).

The results highlight the substantial contribution of BCS to the differences in live weight of the ewe. A linear relationship suggests that, for a given breed type, a single incremental liveweight change across the entire BCS range can be applied. Thus, as a ewe ages, a greater liveweight change is required to alter BCS by one unit, which translates into greater energy requirements in order to make the change, which could have nutritional ramifications (Freer et al., 2007; Morel et al., 2016). The findings also point to the possibility of predicting BCS from live weight and vice versa using a linear regression model. If so then when predicting any of the two variables above, consideration of factors such as age group, stage of the annual production cycle and pregnancy-rank is required and therefore, different prediction equations may be needed.

10.2.5 Chapter 7. Predicting Ewe Body Condition Score Using Lifetime Live weight and Liveweight Change, and Previous Body Condition Score Record

This study aimed to investigate the possibility of using lifetime live weight and liveweight change and previous body condition score to predict current body condition score in Romney ewes. It was found that the equations combining live weight, liveweight change and previous BCS (combined models) explained more variability in BCS (39.8%) and had less prediction error (i.e. 10 to 12%) than equations based on liveweight alone (liveweight alone models). However, a significant portion of the variability in BCS remained unaccounted for (39 to 89%) even in the combined models.

The results indicate that a combination of lifetime live weight, liveweight change and previous body condition score improved body condition score prediction. Given the greater proportion of unexplained variability in BCS, the procedures found in this study, may overestimate or underestimate measures by 0.23 to 0.32 BCS and thus, should be used cautiously. The findings do still suggest that this BCS prediction error could be reduced if key variables affecting the relationship between BCS and live weight are accounted for. This would benefit farmers by allowing for targeted nutritional management of individual animals to maximize overall flock productivity.

10.2.6 Chapter 8. Application of machine learning algorithms to predict body condition score from live weight records of mature Romney ewes

This study utilized selected machine learning (ML) classification algorithms to explore the possibility of predicting BCS of ewes at 43 to 54 months of age on a 3-point scale (1.0−2.0, 2.5−3.5; >3.5) from current and previous live weights. It was found that greater BCS prediction accuracies were achievable (> 85%) across all stages of the annual ewe production cycle using black-box ML (such as boosted trees classification: XGB) than the more conventional models (Ordinal, multinomial regression) or the Classification and Regression Tree. Additionally, all models had balanced specificity and sensitivity (authenticity). For the first time, the study devised a unified indicator for model prediction performance combining several accuracy and authenticity on a single platform.

Combined the results suggest that with class balance, ewe BCS can be predicted with great accuracy and authenticity from a ewe's current and previous live weight using machine learning algorithms. Further, with more variability in BCS explained, through accounting for key variables affecting the relationship between BCS and LW, the accuracy could be ameliorated, and this warrants research. These algorithms if trained on a large representative dataset, could be incorporated into weighing systems to easily and quickly give farmers accurate ewe BCS prediction/categorization without the need for hands-on burden.

10.2.7 *Chapter 9: Predicting ewe body condition score using adjusted live weight for conceptus and fleece weight, height at withers and previous body condition score record*

This study investigated the possibility of improving the accuracy of ewe BCS prediction by using a linear combination of adjusted live weights (correcting for conceptus live weight and fleece weight) and height at withers. It was found that using adjusted live weights and height at withers in addition to previous BCS did not improve the current BCS prediction accuracy. In addition, a considerable portion of unexplained variability in BCS remained.

The results indicate that using adjusted LW or adding height at withers data in a linear combination offered no added advantage to current BCS prediction. Given the great prediction error and proportions variability in BCS still unexplained, it appears that collecting additional production characteristics data by farmers to help account for conceptus and fleece weight would not be useful. However, it is possible that if machine learning models in Chapter 8 were applied on adjusted live weights, or if other technologies to account for individual variability in BCS were applied, more accurate BCS could be achieved.

10.3 Limitations of the study

The initial ewe lamb liveweight loss study (Chapter 3) utilised 40 ewes, across two groups, with group sizes of 20. As this was a study nested in an already running study, it was not possible to have enough sample space to randomly select ewes for the nested study trial. As a result, the initial live weights for the study ewes were similar with a limited range of live weight. This could have affected the rate of liveweight loss during fasting. Examination of the live weight decay curves, however, showed that there was great variability in the rate of liveweight loss from ewes with comparable initial live weight, suggesting that the limited diversity of initial live weights could have not had effect of the rate of liveweight loss. The observed individual variability in the rate of liveweight loss indicate that there could be innate individual differences altering the live weight decay curves. Such individual differences if accounted for may improve the accuracy of liveweight loss prediction and warrant further research. Further, in Chapter 3, two types of diets were evaluated (Ryegrass-based and clover-based swards). By the time this thesis was written, the results from analysis of feed that these ewes were consuming had not been obtained. Therefore, the explanations given for the differences in liveweight loss between Ryegrass and Clover could not be bolstered through evidence i.e. attributed to what they were fed (Dy matter, nutrients, minerals) and may not be conclusive. This warrants investigation. In the next set of studies (Chapter 4 and Chapter 5), the predominant Ryegrass-based diet was carried forward to investigate the effect of herbage availability (allowance), season and physiological state of a ewe using group sizes of 25-30 animals generated with a power of 0.9. In these studies, quantity and quality of feed offered was measured. These studies have given more conclusive results and showed that differences in Dry matter content of feed directly influenced the rate of ewe liveweight loss, an aspect that was not apparently substantiated in

Chapter 3. These studies (Chapter 4 and Chapter 5), used different ewe ages (i.e. ewe lambs in Chapter 4 and mixed-aged ewes in Chapter 5), subjected to different herbage availability levels (ewe lambs: 700-900 Kg DM/ha; 110-130, >1400; ewes: 900-1100 Kg DM/ha, >1400) making it not possible to directly compare the performance of both ewe groups. This was further confounded by the fact that studies were not conducted at the same time and on ewes in the same physiological state. Different times of the year come with different challenges such as scarcity of feed and water and physiological ewe stage. In Chapter 5, during both the calibration and validation stages of the non-pregnant ewe study, there were challenges of limited green grass. As a result, one of the farms where validation was conducted had barely any green grass (DM > 89%), while the calibration equations were developed using lower DM grass (DM > 30%). This likely greatly affected the accuracy of the live weight correction equations. In addition, in all liveweight loss studies, (Chapters 3, 4, and 5), ewes were removed from feed and water in the morning (9:00 – 10:00 am), weighed on arrival at the weighing facility and then hourly. Time of day relative length of day can affect the gut-fill and hence the rate of liveweight loss. It is, therefore, possible that if ewes were removed from pasture at any other time of day, this would affect the accuracy of the developed correction equations and may warrant further studies.

In Chapters 4 and 5, herbage availability target ranges were maintained. The current studies were conducted on commercial farms, which limited control over the preparation and management of herbage availability levels in the grazing paddocks. Ideally, each paddock should have been maintained within a narrow range (±50 kg DM/ha) which equates to the error associated with the herbage availability estimating equation (Hodgson et al., 1999). Observance of the herbage availability target ranges was limited by the number and area of available paddocks and livestock available to control the herbage availability. Using multiple farms with similar conditions and management practices to provide more grazing area, and a mechanical mower to keep herbage within the desired availability levels would reduce the effects of the above challenges. To reduce the effect of individual paddocks, new sets of paddocks were used during replication of the study. Further, treatment groups were replicated across both seasons, farms, years, and studies, providing confidence in results obtained. It is, therefore, unlikely that lack of strict observance of the herbage availability ranges could have affected the outcomes of the current studies.

Weight measurements were carried out using static weighing systems on two different facilities i.e. open space and a roofed shelter. Therefore, this required the physical collection of ewes from their paddocks, mustering and fasting for a specific period. Differences in wind exposure in these weighing facilities can distort the accuracy of weight measurements. Further, in static facilities one may need to drive the novice ewes through the weighing crate. This may make it time consuming. It was assumed that the effect of wind pressure onto the loading bars was inconsequential. Ideally, all ewe weighing should have been conducted under the same farm conditions. This could not be observed as available farms had different weighing facility types which is typical of the commercial farms in New Zealand. To reduce the effect of abrupt wind pressure especially on wet days, most weighing was conducted on dry days or in roofed shelters. This could be overcome by standardizing the weighing protocol to ensure similar weighing facilities are used. However, the effect of using different weighing facility types was minimized by replicating the studies across these farms with different weighing facilities.

Day of pregnancy can affect the rate of live weight (Burnham et al., 2009) and the rate of conceptus growth (Kleemann and Walker, 2005; Kenyon et al., 2008) affecting the measurement of live weight, and thus, their accuracy is of utmost importance. For a large flock under extensive management, it is not possible to record the mating date of an individual ewe to subsequently estimate its day in pregnancy. In the current studies (Chapter 6, 7, 8, 9), individual day of pregnancy were estimated using the midpoint of a 17-day breeding period as the reference day. This standardized and increased the certainty in estimation of the day of pregnancy and consequently the measurements associated.

Due to the subjective nature of BCS, reliability within and between assessors can be a stumbling block to the potential use and effectiveness of this technique (Kenyon et al., 2014). This is especially true when different inexperienced assessors are used to determine BCS measurements. In the BCS profiling (Chapter 6) and prediction studies (Chapter 7, 8), initially, BCS measurements were determined by two experienced assessors (one for the first 5 years and one for the final year of study) and later (Chapter 9, for all 3 years) by single assessor. For experienced assessors, reliability levels of up to 90% have been reported. It is, therefore, unlikely that the reliability of the BCS data used in the current study was greatly affected.

BCS data (on 1.0−5.0 scale) from the current study was highly imbalanced. This imbalance was most apparent in the extreme BCS values/classes (1.0−2.0 and 4.0−5.0). When linear regression models (as in Chapters 7 and 9) are applied on unbalanced data (BCS as continuous variable), the imbalance is of little consequence. However, this would likely affect the accuracy of results in a categorical BCS prediction study (as in Chapter 8). To reduce the effect of class imbalance, a less imbalanced three-point BCS scale (1.0−2.0, 2.5−3.5, >3.5) was devised guided by literature (Kenyon et al., 2014) for categorical BCS prediction (Chapter 8). Further, resampling techniques (SMOTE) were employed to generate synthetic sample sizes representative of the original data structure (Chawla et al., 2002), thereby restoring the assumption of proportional odds and providing confidence in the findings. Therefore, the process of developing the current strategies for BCS prediction was less affected by the effects of class imbalance.

Effects of gut-fill fluctuation and fleece weight on live weight measurement can be significant. These can confound the relationship between LW and other measurements. In Chapters 6, 7 and 8, only LW and BCS data were collected. Consequently, unadjusted LW and LW change were used in the analysis which could potentially have affected the accuracy of the models relating LW to BCS. This was solved in Chapter 9 where additional production data were collected and used to adjust for effects of conceptus weight and fleece weight. The results in

Chapter 9 suggest that adjusting for live weight had no impact on the overall accuracy the BCS prediction. Therefore, it is unlikely that this lack of adjusting for conceptus and fleece weight affected the findings in the previous Chapters. However, it was not possible to apply the developed LW correcting equations in Chapter 9 as both studies were running concurrently.

In Chapter 9, standard Gompertz's function for adjusting foetal weight and Cottle for fleece weight adjustment were used to correct for LW. Although such equations have been used for quite some time, it is not known how much error (noise) they are likely to introduce rather than remove. It is, therefore, likely that these equations could have introduced more noise than removed it and this may warrant further examination. Additionally, in the prediction of ewe BCS studies, a great part of variation was not explained by the production traits. This variation is largely attributed to individual ewe differences. Strategies to account for these individual differences including accounting for individual frame sizes should be investigated.

10.4 Next steps in research required

Current live weight and BCS measurement improvement strategies have been developed using a series of studies conducted primarily in Manawatū-Whanganui and Wairarapa regions in New Zealand. In addition, the current study utilized one breed of sheep (Romney). It is not known if these live weights and BCS measurement improvement strategies are appropriate to other regions of New Zealand for other breeds. Therefore, further studies to evaluate the feasibility of these live weight and BCS measurement improvement strategies to improve ewe performance assessment in other regions of New Zealand and breeds of sheep would likely be of benefit. The current study utilized data from institutional research farms. It is not known if the developed strategies would be appropriate for privately-run farms and may warrant investigation.

Current strategies (Chapters 3, 4 and 5) accounted effects of two herbage types and availability, and three seasons in a limited ewe physiological state range and no data for male sheep. It is not known to what extent these strategies are appropriate for all herbage types, seasons, ewe physiological states, and for male sheep. Further, the results of this thesis indicate that there was more variability in liveweight loss rates that were not accounted for. It is possible that increases in accuracies greater than what is reported above would be achievable. This variability in liveweight loss rates was attributed to the wide herbage availability target ranges used in this study. Ideally, maintaining these herbage availability target ranges within ± 50 kg DM/ha should increase the accuracy of "without delay" ewe live weight. Sheep feed intake is expected to increase with herbage availability up to 1400 kg DM/ha beyond which it is expected to remain constant (Morris and Kenyon, 2004). It is possible that variations in liveweight loss rates above 1400 kg DM/ha, could be attributed to factors beyond herbage availability which warrant further research. Additionally, the strategies developed in this study were largely based on ryegrass-based herbages. Herbages can vary in chemical composition (Cranston et al., 2015; Ekanayake et al., 2019) and digesta kinetics (Moyo and Nsahlai, 2018). It is not known to what extent these interventions could be applicable to other herbage types (e.g. plantain, native shrubs). Further, the study animals were all weighed at the same time of day (i.e. 9:00 to 10:00am). The time of weight measurement in relation to length of day can influence the intake, gut-fill and digesta kinetics. It is, therefore, worthwhile investigating the effect of time day sheep were removed from the pastures. In addition, the current strategies to correct LW, utilized a static weighing system where ewes were fasted. Elsewhere, Walk-over weighing systems are used by farmers. A comparative cost benefit study of using the developed correction equations, providing a feed supplement at a static weighing facility and/or using a walk-over weighing system will be vital.

Current strategies (Chapters 7 and 9) further, examined the possibility of predicting BCS on a full scale (1.0−5.0) from a ewe's production characteristics using linear regression. A significant proportion of the variability in BCS remained unexplained by the models. It implies that factors beyond the production characteristics could be responsible for the unexplained variance and this warrants research. The great variability in unexplained ewe BCS could be attributed in individual differences. Therefore, further studies using data aimed such as image analysis (Computed Tomography: CT scans, dual-energy X-ray absorptiometry: DXA), DNA profiles and automated metabolic profiles, to account for individual animal variability may be warranted. Further, the use of machine learning techniques (Chapter 8) to predict BCS from a ewe's unadjusted live weights (unadjusted for conceptus and fleece weight) achieved accuracies greater than 85% on a narrow 3-point scale. This machine learning tool presents an opportunity for screening out the thin (BCS < 2.5) and obese (BCS > 3.5) ewes while preserving those in optimal productivity range (2.5 < BCS <3.5). It is possible that accuracies beyond and above 85% could be achieved if adjusted live weight and liveweight changes were used in the BCS prediction models. It, therefore, warrants further research to estimate the feasibility of increasing the accuracy of machine learning BCS prediction models when additional ewe characteristics are supplied.

The current strategies used a discrete and rigid BCS scale based on a subjective method. A BCS scale of 1.0-5.0 with increments of 0.5 points was applied in the present studies. Elsewhere such as in Australia, increments of 0.25 have been used. It is possible that by increasing the length of the scale to include more points, more accurate BCS predictions can be achieved. Therefore, future studies should investigate the impact of using a longer scale (such as BCS: 1.0 – 10.0) or a scale with more increments (BCS: 1.0 - 5.0; 0.25 points). Further, given the limitations of predicting BCS, itself a predictor of body composition. It would be worthwhile investigating how accurately live weights and other predictors would predict total body fat and muscle weights, or proportions given they are more objective and continuous variables. The first step in these types of studies would require animals to be euthanized and/or tools such as CT scans.

The supply of additional data to adjust LW for conceptus weight and fleece weight in the study did not appear to improve the prediction of BCS. In these strategies we used standard equations whose error rates and reliability are unknown. Therefore, could have introduced more noise than removed it. Therefore, future studies should investigate the reliability and extent of application of these standard equations or even develop customised equations for the study population.

10.5 Practical implications and recommendations

In the southern hemisphere, sheep production is mainly extensive with large flock sizes where management is based on average flock performance (Kenyon et al., 2014; Brown et al., 2015). Ewe performance is usually assessed using live weights and body condition score. Assessments that are based on a group rather than an individual's performance can be misleading and may impede decision making and consequently proper allocation of resources. The results of this thesis supported the possibility of recording accurate live weights using a set of correcting equations and suggested that handsfree BCS measurement for improved ewe performance assessment was possible. Collection of accurate data has been shown to improve decision making regarding the management of ewes and consequently increased their performance (Young et al., 2004; Curnow et al., 2011). It is possible that software incorporating live weight corrections and BCS estimations could be developed and made available in already existing weight electronic systems. These live weight correcting capabilities would adjust data based on herbage offered, ewe age, ewe physiological state, and season, when computing the final "without delay" live weight. The farmer would be required to record the time when ewes were removed from pasture and enter this into the automatic weighing system. These systems already have individual time stamp capability for individual weights. For ewe BCS, the predictions would be based on their individual live weight and BCS records collected routinely. The current study has shown that collection of more ewe production data (Chapter 9: conceptus weight, fleece weight, height at withers) would be of no benefit, consequently, farmers would have no extra burden of recording these additional variables. Therefore, it should be possible to have both corrected live weight and BCS recorded automatically by the weighing system if correct management information (age, physiological status, herbage type and availability) is pre-determined. The cost benefit of this would require further evaluation. If accurate equations can be developed to predict without delay live weight and current BCS, then farmers would be able to make more informed decisions.

10.6 Overall summary and conclusions

A series of studies have been undertaken to determine the effects of various factors on the rate of ewe liveweight loss during weighing, the possibility of correcting for live weights and predicting BCS using a ewe's characteristics. Briefly, the studies have led to the following conclusions:

- The rate of ewe liveweight loss depends on type (quality) and availability (quantity) of herbage, season of year and the physiological state of a ewe. Within physiological state, stage of pregnancy influences the rate of liveweight loss but not pregnancy-rank.
- Applying correction equations reduces the error associated with delayed weights and improves the accuracy of "without delay" live weight estimates.
- The relationship between ewe live weight and body condition score is linear and influenced by age, stage of the annual production cycle and pregnancy-rank.
- Utilizing live weight, liveweight change and previous BCS record to predict a ewe's current BCS using a general linear model improves the prediction of BCS but does not explain much of the variability in BCS.

- Machine learning improves ewe BCS prediction from a ewe's live weight records.
- Utilizing additional ewe information such as stage of gestation, pregnancy-rank and lambing records, and fleece weight to correct ewe live weights for conceptus and fleece weight and height at withers in addition to liveweight change and previous BCS does not improve the prediction of BCS.

Overall, the findings from the current study suggest that measurement of both LW and BCS can be improved. Standardized feeding prior to ewe weighing and use of live weight correction equations improved the accuracy of delayed live weight estimates relative to "without delay" weights. Further, LW and BCS are linearly related and this relationship is influenced by ewe age, stage of the annual ewe production cycle and pregnancy-rank. If such factors affecting LW and BCS can be accounted for, it is possible to exploit this relationship between BCS and LW to predict BCS. These are important findings which will provide useful platform for future studies aiming to manipulate weighing and BCS protocols and systems to improve sheep management.

References

Abdel-Mageed, I. (2009). Body condition scoring of local Ossimi ewes at mating and its impact on fertility and prolificacy. *Egyptian Journal of Sheep and Goat Sciences* 4, 37-44.

Adalsteinsson, S. (1979). The independent effects of live weight and body condition on fecundity and productivity of Icelandic ewes. *Animal Science* 28, 13-23.

Agresti, A. (1999). Modelling ordered categorical data: recent advances and future challenges. *Statistics in medicine* 18, 2191-2207.

Agresti, A., and Kateri, M. (2011). Categorical Data Analysis. *In* "International Encyclopedia of Statistical Science" (M. Lovric, ed.), pp. 206-208. Springer Berlin Heidelberg, Berlin, Heidelberg.

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine* 18, 91-93.

Aktaş, A. H., Dursun, Ş., Doğan, Ş., Kiyma, Z., Demirci, U., and Halıcı, I. (2015). Effects of ewe live weight and age on reproductive performance, lamb growth, and survival in Central Anatolian Merino sheep. *Archives Animal Breeding* 58, 451-459.

Al-Ramamneh, D., Riek, A., and Gerken, M. (2012). Effect of water restriction on drinking behaviour and water intake in German black-head mutton sheep and Boer goats. *Animal* 6, 173-178.

Al-Sabbagh, T. A., Swanson, L. V., and Thompson, J. M. (1995). The effect of ewe body condition at lambing on colostral immunoglobulin G concentration and lamb performance. *Journal of Animal Science* 73, 2860-2864.

Alamer, M. (2009). Effect of water restriction on lactation performance of Aardi goats under heat stress conditions. *Small Ruminant Research* 84, 76-81.

Alberts, B., and Abdi, H. (2007). Bonferroni and Šidák Corrections for Multiple Comparisons. *In* "Encyclopedia of Measurement and Statistics" (N. Salkind, ed.), Vol. 27, pp. 103-107. Sage., Thousand Oaks, CA.

Alexander, D. L. J., Tropsha, A., and Winkler, D. A. (2015). Beware of R(2): Simple, Unambiguous Assessment of the Prediction Accuracy of QSAR and QSPR Models. *Journal of Chemical Information and Modeling* 55, 1316-1322.

Alexopoulos, E. C. (2010). Introduction to multivariate regression analysis. *Hippokratia* 14, 23-28.

Aliyari, D., Moeini, M. M., Shahir, M. H., and Sirjani, M. A. (2012). Effect of body condition score, live weight and age on reproductive performance of Afshari ewes. *Asian Journal of Animal and Veterinary Advances* 7, 904-909.

Allison AJ, K. R. W. (1978). "Some effects of liveweight and breed of ewe on fertility and fecundity," The society of Sheep and Beef Cattle Veterinarians of the New Zealand Veterinary Association.

Alwash, A. H., and Thomas, P. C. (1971). The effect of the physical form of the diet and the level of feeding on the digestion of dried grass by sheep. *Journal of the Science of Food and Agriculture* 22, 611-615.

Anusha, V. (2016). Utility of body condition score (BCS) system for the assessment of physical traits and performance of Nellore (brown) sheep. . M.V.Sc, Sri Venkateswara Veterinary University Tirupati, India, pp. 1-130.

Archer, K. J., and Williams, A. A. (2012). L 1 penalized continuation ratio models for ordinal response prediction using high-dimensional datasets. *Statistics in Medicine* 31, 1464-1474.

Atti, N., Thériez, M., and Abdennebi, L. (2001). Relationship between ewe body condition at mating and reproductive performance in the fat-tailed Barbarine breed. *Animal Research* 50, 135-144.

Azzaro, G., Caccamo, M., Ferguson, J. D., Battiato, S., Farinella, G. M., Guarnera, G. C., Puglisi, G., Petriglieri, R., and Licitra, G. (2011). Objective estimation of body condition score by modeling cow body shape from digital images. *Journal of Dairy Science* 94, 2126-2137.

Baker, A. L., Phillips, R. W., and Black, W. H. (1947). The relative accuracy of one-day and three-day weaning weights of calves. *Journal of animal science* 6, 56-59.

Bakoev, S., Getmantseva, L., Kolosova, M., Kostyunina, O., Chartier, D. R., and Tatarinova, T. V. (2020). PigLeg: prediction of swine phenotype using machine learning. *PeerJ* 8, e8764-e8764.

Barry, T. N., Wilson, P. R., and Kemp, P. D. (1999). Forage grazing systems to produce venison according to market signals. *Proceedings of the New Zealand Society of Animal Production* 59, 140-144.

Bartocci, S., Amici, A., Verna, M., Terramoccia, S., and Martillotti, F. (1997). Solid and fluid passage rate in buffalo, cattle and sheep fed diets with different forage to concentrate ratios. *Livestock Production Science* 52, 201-208.

Bean, H. W. (1946). Single Weight Versus a Three-Day-Average Weight for Swine. *Journal of Animal Science* 5, 30-35.

Bean, H. W. (1948). Single weight versus a three-day average weight for sheep. *Journal of animal science* 7, 50-54.

Behrendt, R., Hocking Edwards, J. E., Gordon, D., Hyder, M., Kelly, M., Cameron, F., Byron, J., Raeside, M., Kearney, G., and Thompson, A. N. (2019). Offering maternal composite ewes higher levels of nutrition from mid-pregnancy to lambing results in predictable increases in birthweight, survival and weaning weight of their lambs. *Animal Production Science* 59, -.

Bell, M. J., Maak, M., Sorley, M., and Proud, R. (2018). Comparison of Methods for Monitoring the Body Condition of Dairy Cows. *Frontiers in Sustainable Food Systems* 2.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J.-M., and McBratney, A. (2010). Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC Trends in Analytical Chemistry* 29, 1073-1081.

Bernabucci, U., Bani, P., Ronchi, B., Lacetera, N., and Nardone, A. (1999). Influence of Short- and Long-Term Exposure to a Hot Environment on Rumen Passage Rate and Diet Digestibility by Friesian Heifers1. *Journal of Dairy Science* 82, 967-973.

Berry, P. D., Buckley, F., and Dillon, P. (2007). Body condition score and live-weight effects on milk production in Irish Holstein-Friesian dairy cows. *Animal* 1, 1351-1359.

Besier, R. B., and Hopkins, D. (1989). Farmers' estimations of sheep weights to calculate drench dose. *Journal of the Department of Agriculture, Western Australia, Series 4* 30, 120-121.

Bewley, J. M., Peacock, A. M., Lewis, O., Boyce, R. E., Roberts, D. J., Coffey, M. P., Kenyon, S. J., and Schutz, M. M. (2008). Potential for Estimation of Body Condition Scores in Dairy Cattle from Digital Images. *Journal of Dairy Science* 91, 3439-3453.

Bich, W., Cox, M. G., Dybkaer, R., Elster, C., Estler, W. T., Hibbert, B., Imai, H., Kool, W., Michotte, C., and Nielsen, L. (2012). Revision of the 'Guide to the expression of uncertainty in measurement'. *Metrologia* 49, 702.

Bishop, P. A., and Herron, R. L. (2015). Use and misuse of the likert item responses and other ordinal measures. *International journal of exercise science* 8, 297.

Blaikie, N. (2003). "Analyzing quantitative data: From description to explanation," Sage.

Böhning, D. (1992). Multinomial logistic regression algorithm. *Annals of the Institute of Statistical Mathematics* 44, 197-200.

Botchkarev, A. (2019). Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology. *Interdisciplinary Journal of Information, Knowledge, and Management* 14, 45-79.

Box, L. A., Edwards, G. R., and Bryant, R. H. (2018). In sacco digestion kinetics of plantain and ryegrass-white clover harvested in the morning and afternoon. *New Zealand Journal of Animal Science and Production* 78, 34-39.

Branco, P., Ribeiro, R. P., and Torgo, L. (2016). UBL: an R package for utility-based learning. *arXiv preprint arXiv:1604.08079*.

Breiman, L. (1996). Bagging predictors. *Machine learning* 24, 123-140.

Breiman, L. (1998). Arcing classifier (with discussion and a rejoinder by the author). *The annals of statistics* 26, 801-849.

Breusch, T. S., and Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica* 47, 1287-1294.

Broring, N., Wilton, J., and Colucci, P. (2003). Body condition score and its relationship to ultrasound backfat measurements in beef cows. *Canadian journal of animal science* 83, 593-596.

Brosh, A., Shkolnik, A., and Choshniak, I. (1986). Metabolic effects of infrequent drinking and low-quality feed on Bedouin goats. *Ecology* 67, 1086-1090.

Brown, D. J., Ball, A. J., and Huisman, A. E. (2005). The influence of ewe weight at mating on lamb performance and reproduction of the ewe. *Proceedings of the Association for the Advancement of Animal Breeding and Genetics* 16, 306-309.

Brown, D. J., Savage, D. B., and Hinch, G. N. (2014). Repeatability and frequency of in-paddock sheep walk-over weights: implications for individual animal management. *Animal Production Science* 54, 207-213.

Brown, D. J., Savage, D. B., Hinch, G. N., and Hatcher, S. (2015). Monitoring liveweight in sheep is a valuable management strategy: a review of available technologies. *Animal Production Science* 55, 427-436.

Burke, J. A., Nuthall, P. L., and McKinnon, A. E. (2004). "An analysis of the feasibility of using image processing to estimate the live weight of sheep. FHMG research report," Rep. No. 1174-8796. Lincoln University,Canterbury, New Zealand.

Burnham, D. L., Morel, P. C. H., Kenyon, P. R., Morris, S. T., and Stafford, K. J. (2009). The effect of fasting on live weight in hogget wethers and pregnant mature ewes: Brief Communication. *Proceedings of the New Zealand Society of Animal Production* 69, 112-114.

Cain, K. C., Kronmal, R. A., and Kosinski, A. S. (1992). Analysing the relationship between change in a risk factor and risk of disease. *Statistics in medicine* 11, 783-797.

Cake, M. A., Gardner, G. E., Boyce, M. D., Loader, D., and Pethick, D. W. (2006). Forelimb bone growth and mineral maturation as potential indices of skeletal maturity in sheep. *Australian Journal of Agricultural Research* 57, 699-706.

Calavas, D., Sulpice, P., Lepetitcolin, E., and Bugnard, F. (1998). Assessing the accuracy of the practice of a method of scoring the body condition of ewes within a professional framework. *Veterinary research* 29, 129-138.

Casellas, J., Caja, G., Piedrafita, J., and Such, X. (2007). Survival analysis from birth to slaughter of Ripollesa lambs under semi-intensive management1. *Journal of Animal Science* 85, 512-517.

Chan, Y. (2003). Biostatistics 104: correlational analysis. *Singapore Medical Journal* 44, 614-9.

Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research* 16, 321-357.

Chay-Canul, A. J., Garcia-Herrera, R., Meza-Villalvazo, V. M., Gomez-Vazquez, A., Cruz-Hernandez, A., Magaña-Monforte, J. G., and Ku-Vera, J. C. (2016). Body fat reserves and their relationship to ultrasound back fat measurements in Pelibuey ewes. *Ecosistemas y Recursos Agropecuarios* 3, 407-413.

Chay-Canul, A. J., Garcia-Herrera, R. A., Robertos, N., Macias-Cruz, U., Vicente-Pérez, R., and Meza-Villalvazo, V. M. (2019). Relationship between body condition score and subcutaneous fat and muscle area measured by ultrasound in Pelibuey ewes. *Emirates Journal of Food and Agriculture* 31 53-58.

Chen, L.-F., Liao, H.-Y. M., Ko, M.-T., Lin, J.-C., and Yu, G.-J. (2000). A new LDA-based face recognition system which can solve the small sample size problem. *Pattern recognition* 33, 1713-1726.

Chen, T., and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *In* "Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining" (B. Krishnapuram, ed.), pp. 785-794, San Francisco, USA.

Christopherson, R. J. (1976). Effects of prolonged cold and the outdoor winter environment on apparent digestibility in sheep and cattle. *Canadian Journal of Animal Science* 56, 201-212.

Coates, D. B., and Penning, P. (2000a). "Measuring animal performance," pp. 1-353-402.
Coates, D. B., and Penning, P. (2000b). Measuring animal performance. *In* "Field and laboratory methods for grassland and animal production research" (L. t. a. R. Jones, ed.), pp. 353-402. CABI Publishing, Wallingford, UK.
Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20, 37-46.
Collins, C. T., and Atwood, J. L. (1981). Weighing with spring balances: Biases and errors. *North American Bird Bander* 6.
Conover, W. J. (1973). On Methods of Handling Ties in the Wilcoxon Signed-Rank Test. *Journal of the American Statistical Association* 68, 985-988.
Cook, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics* 19, 15-18.
Corner-Thomas, R. A., Hickson, R. E., Morris, S. T., Back, P. J., Ridler, A. L., Stafford, K. J., and Kenyon, P. R. (2015b). Effects of body condition score and nutrition in lactation on twin-bearing ewe and lamb performance to weaning. *New Zealand Journal of Agricultural Research* 58, 156-169.
Corner-Thomas, R. A., Kenyon, P. R., Morris, S. T., Ridler, A. L., Hickson, R. E., Greer, A. W., Logan, C. M., and Blair, H. T. (2016). Brief communication: The use of farm-management tools by New Zealand sheep farmers: changes with time. *Proceedings of the New Zealand Society of Animal Production* 76, 78-80.
Corner-Thomas, R. A., Ridler, A. L., Morris, S. T., and Kenyon, P. R. (2015a). Ewe lamb live weight and body condition scores affect reproductive rates in commercial flocks. *New Zealand Journal of Agricultural Research* 58, 26-34.
Corson, D., Waghorn, G., Ulyatt, M., and Lee, J. (1999). NIRS: Forage analysis and livestock feeding. *Proceedings of the New Zealand Grassland Association*, 127-132.
Cottle, D., and Pacheco, D. (2017). Prediction of fleece insulation after shearing and its impact on maintenance energy requirements of Romney sheep. *Small Ruminant Research* 157, 14-22.
Craig, R. L. (1982). Breeding from Romney ewe hoggets in the Waihora group breeding scheme. *New Zealand Agricultural Science* 16, 101-104.
Crampton, E. W., and Jackson, I. R. C. (1944). Seasonal Variation in Chemical Composition of Pasture Herbage and the Relation to Its Digestibility by Steers and Sheep. (Pasture Studies—XXVI). *Journal of Animal Science* 3, 333-339.
Cranston, L., Kenyon, P., Morris, S., and Kemp, P. (2015). A review of the use of chicory, plantain, red clover and white clover in a sward mix for increased sheep and beef production. *Journal of New Zealand Grasslands*, 89-94.
Cranston, L., Ridler A, A, G., and P., K. (2017). Pasture measurement. *In* "Livestock Production in New Zealand" (S. K, ed.), pp. 88–96. Massey University Press, Auckland New zealand.
Curnow, M., Oldham, C. M., Behrendt, R., Gordon, D. J., Hyder, M. W., Rose, I. J., Whale, J. W., Young, J. M., and Thompson, A. N. (2011). Successful adoption of new guidelines for the nutritional management of ewes is dependent on the development of appropriate tools and information. *Animal Production Science* 51, 851-856.

Daniel, G. G. (2013). Artificial Neural Network. *In* "Encyclopedia of Sciences and Religions" (A. L. C. Runehov and L. Oviedo, eds.), pp. 143-143. Springer Netherlands, Dordrecht.

Davis, G. H., McEwan, J. C., Howarth, M. V., Kelly, R. W., and Paterson, A. T. (1987). Reproductive performance of progeny of prolific Romney, Coopworth and Perendale sheep. *In* "Proc. NZ Soc. Anim. Prod.", Vol. 47, pp. 89-91.

Dhakal, C. P. (2018). Multiple Regression Model Fitted for Rice Production Forecasting in Nepal: A Case of Time Series Data. *Nepalese Journal of Statistics* 2, 89-98.

Dietterich, T. G. (2000). Ensemble methods in machine learning. *In* "International workshop on multiple classifier systems", pp. 1-15. Springer.

Dinga, R., Penninx, B. W., Veltman, D. J., Schmaal, L., and Marquand, A. F. (2019). Beyond accuracy: measures for assessing machine learning models, pitfalls and guidelines. *bioRxiv*, 743138.

Dodds, K. G., and Everett-Hincks, J. M. (2008). Management of maternal-offspring behavior to improve lamb survival in easy care sheep systems1,2. *Journal of Animal Science* 86, E259-E270.

Dosne, A.-G., Bergstrand, M., and Karlsson, M. O. (2016). A strategy for residual error modeling incorporating scedasticity of variance and distribution shape. *Journal of Pharmacokinetics and Pharmacodynamics* 43, 137-151.

Dowman, M. G., and Collins, F. C. (1982). The use of enzymes to predict the digestibility of animal feeds. *Journal of the Science of Food and Agriculture* 33, 689-696.

Drosg, M. (2009). "Dealing with uncertainties: a guide to error analysis," 2nd/Ed. Springer Science & Business Media.

Durgesh, K. S., and Lekha, B. (2010). Data classification using support vector machine. *Journal of theoretical and applied information technology* 12, 1-7.

Dyrmundsson, O. R. (1973). Puberty and early reproductive performance in sheep. 1. Ewe lambs. *Animal Breeding Abstracts* 41, 237-289.

Ebrahimi, M., Mohammadi-Dehcheshmeh, M., Ebrahimie, E., and Petrovski, K. R. (2019). Comprehensive analysis of machine learning models for prediction of sub-clinical mastitis: Deep learning and gradient-boosted trees outperform other models. *Computers in biology and medicine* 114, 103456.

Edey, T. N. (1968). Body weight and ovulation rate in sheep. *Proceedings of the Australian Society of Animal Production* 7, 188-191.

Efron, B. (1983). Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation. *Journal of the American Statistical Association* 78, 316-331.

Efron, B. (2014). Estimation and Accuracy after Model Selection. *Journal of the American Statistical Association* 109, 991-1007.

Ekanayake, W. E. M. L. J., Corner-Thomas, R. A., Cranston, L. M., Hickson, R. E., Kenyon, P. R., and Morris, S. T. (2019). Characterisation of the nutritional composition of plant components of a herb-clover mix during November to May in New Zealand. *New Zealand Journal of Animal Science and Production* 79, 162-167.

Elliott, K. H., Bigham, M. L., Sumner, R. M. W., and Dalton, D. C. (1978). Wool production of yearling ewes of different breeds on hill country. *New Zealand Journal of Agricultural Research* 21, 179-186.

Elwood, M. (2017). Critical appraisal of epidemiological studies and clinical trials. pp. 483. Oxford University Press, Oxford, United Kingdom.

Evans, D. G. (1978). The interpretation and analysis of subjective body condition scores. *Animal Science* 26, 119-125.

Everitt, G. C. (1962). On the assessment of body composition in live sheep and cattle. *Proceedings of the Australian Society of Animal Production conference* 4, 79-94.

Ferguson, M. B., Thompson, A. N., Gordon, D. J., Hyder, M. W., Kearney, G. A., Oldham, C. M., and Paganoni, B. L. (2011). The wool production and reproduction of Merino ewes can be predicted from changes in liveweight during pregnancy and lactation. *Animal Production Science* 51, 763-775.

Ferra, J. d. C., Cieslak, S., Sartori Filho, R., McManus, C., Martins, C. F., and Sereno, J. R. B. (2010). Weight and age at puberty and their correlations with morphometric measurements in crossbred breed Suffolk ewe lambs. *Revista Brasileira de Zootecnia* 39, 134-141.

Fisher, D. H. (1987). Knowledge acquisition via incremental conceptual clustering. *Machine learning* 2, 139-172.

Fleiss, J. L., Levin, B., and Paik, M. C. (1981). The measurement of interrater agreement. *Statistical methods for rates and proportions* 2, 22-23.

Foot, J. Z., and Russel, A. J. F. (1979). The relationship in ewes between voluntary food intake during pregnancy and forage intake during lactation and after weaning. *Animal Science* 28, 25-39.

Forbes, J. M. (1969). The effect of pregnancy and fatness on the volume of rumen contents in the ewe. *The Journal of Agricultural Science* 72, 119-121.

Forbes, J. M. (1970). Voluntary Food Intake of Pregnant Ewes. *Journal of Animal Science* 31, 1222-1227.

Forcada, F., Abecia, J. A., and Sierra, I. (1992). Seasonal changes in oestrus activity and ovulation rate in Rasa Aragonesa ewes maintained at two different body condition levels. *Small Ruminant Research* 8, 313-324.

Freer, M., Dive, H., and Nolan, J. V. (2007). "Nutrient requirements of domesticated ruminants," CSIRO publishing, Collingwood, Australia, pp. 1-265.

Freer, M., Moore, A. D., and Donnelly, J. R. (1997). GRAZPLAN: Decision support systems for Australian grazing enterprises—II. The animal biology model for feed intake, production and reproduction and the GrazFeed DSS. *Agricultural Systems* 54, 77-126.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software* 33, 1-22.

Frutos, P., Mantecón, A., and Giráldez, F. (1997). Relationship of body condition score and live weight with body composition in mature Churra ewes. *Animal Science* 64, 447–452.

Gabr, A. A., Shalaby, N. A., and Ahmed, M. E. (2016). Effect of ewe born type, growth rate and weight at conception on the ewe subsequent productivity of Rahmani sheep. *Asian J Anim Vet Adv* 11, 732-736.

Galdi, P., and Tagliaferri, R. (2018). Data mining: accuracy and error measures for classification and prediction. *Encyclopedia of Bioinformatics and Computational Biology*, 431-436.

Galwey, J. M., Logan, C. M., and Greer, A. W. (2013). Brief communication: Can weighing multiple times increase the accuracy of live weight recording? *New Zealand Society of Animal Production* 73, 146-148.

Geenty, K. G., Smith, A. J., Dyall, T. R., Lee, G. J., Smith, D., Brewer, H., and Uphill, G. C. (2007). Remote drafting technology for management of pregnant Merino ewes. *In* "Proceedings from the Recent Advances in Animal Nutrition in Australia" (C. P and R. N, eds.), pp. 223-228, University of NewEngland Publishing Unit: Armidale, NSW.

Geisler, P. A., and Fenlon, J. S. (1979). The effects of body weight and its components on lambing performance in some commercial flocks in Britain. *Animal Production* 28, 245-255.

Gibb, M. J., and Treache, T. T. (1980). The effect of ewe body condition at lambing on the performance of ewes and their lambs at pasture. *The Journal of Agricultural Science* 95, 631-640.

Gibb, M. J., and Treacher, T. T. (1982). The effect of body condition and nutrition during late pregnancy on the performance of grazing ewes during lactation. *Animal Science* 34, 123-129.

Ginane, C., and Dumont, B. (2010). Do grazing sheep use species-based categorization to select their diet? *Behavioural Processes* 84, 622-624.

Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and psychological measurement* 55, 377-393.

Golding, K. P., Wilson, E. D., Kemp, P. D., Pain, S. J., Kenyon, P. R., Morris, S. T., and Hutton, P. G. (2011). Mixed herb and legume pasture improves the growth of lambs post-weaning. *Animal Production Science* 51, 717-723.

Gonzalez, R. E., Labuonora, D., and Russel, A. J. E. (1997). The effects of ewe live weight and body condition score around mating on production from four sheep breeds in extensive grazing systems in Uruguay. *Animal Science* 64, 139-145.

Goopy, J. P., Pelster, D. E., Onyango, A., Marshall, K., and Lukuyu, M. (2018). Simple and robust algorithms to estimate liveweight in African smallholder cattle. *Animal Production Science* 58, 1758-1765.

Greene, W. (2003). "Econometric analysis," 5th/Ed. Prentice Hall, Upper Saddle River, New York.

Greenwood, P. L., Hunt, A. S., Hermanson, J. W., and Bell, A. W. (1998). Effects of birth weight and postnatal nutrition on neonatal sheep: I. Body growth and composition, and some aspects of energetic efficiency. *Journal of animal science* 76, 2354-2367.

Greenwood, P. L., Thompson, A. N., and Ford, S. P. (2010). Postnatal Consequences of the Maternal Environment and of Growth During Prenatal Life for Productivity of Ruminants. *In* "Managing the prenatal environment to enhance livestock productivity" (A. W. P L Greenwood, Bell, P E Vercoe and G J Viljoen, ed.). Springer.

Gregorini, P. (2012). Diurnal grazing pattern: its physiological basis and strategic management. *Animal Production Science* 52, 416-430.

Gregorini, P., Gunter, S. A., Beck, P. A., Soder, K. J., and Tamminga, S. (2008). REVIEW: The Interaction of Diurnal Grazing Pattern, Ruminal Metabolism, Nutrient Supply, and Management in Cattle. *The Professional Animal Scientist* 24, 308-318.

Gunn, R. G., and Doney, J. M. (1975). The interaction of nutrition and body condition at mating on ovulation rate and early embryo mortality in Scottish Blackface ewes. *The Journal of Agricultural Science* 85, 465-470.

Gunn, R. G., and Doney, J. M. (1979). Fertility in Cheviot ewes. 1. The effect of body condition at mating on ovulation rate and early embryo mortality in North and South Country Cheviot ewes. *Animal Science* 29, 11-16.

Gunn, R. G., Doney, J. M., and Russel, A. J. F. (1969). Fertility in Scottish Blackface ewes as influenced by nutrition and body condition at mating. *The Journal of Agricultural Science* 73, 289-294.

Gunn, R. G., Doney, J. M., and Russel, A. J. F. (1972). Embryo mortality in Scottish Blackface ewes as influenced by body condition at mating and by post-mating nutrition. *The Journal of Agricultural Science* 79, 19-25.

Gunn, R. G., Maxwell, T. J., Sim, D. A., Jones, J. R., and James, M. E. (1991a). The effect of level of nutrition prior to mating on the reproductive performance of ewes of two welsh breeds in different levels of body condition. *Animal Science* 52, 157-163.

Gunn, R. G., Rhind, S. M., Maxwell, T. J., Sim, D. A., Jones, J. R., and James, M. E. (1988). The effect of sward height and active immunization against androstenedione on reproductive performance of ewes of two Welsh breeds in different body conditions. *Animal Production* 46, 417-426.

Gunn, R. G., Smith, W. F., Senior, A. J., Barthram, E., Sim, D. A., and Hunter, E. A. (1991). Pre-mating herbage intake and the reproductive performance of north Country Cheviot ewes in different levels of body condition. *Animal Science* 52, 149-156.

Gunn, S. R. (1998). Support vector machines for classification and regression. *ISIS technical report* 14, 5-16.

Haaland, G. L., and Tyrrell, H. F. (1982). Effects of Limestone and Sodium Bicarbonate Buffers on Rumen Measurements and Rate of Passage in Cattle. *Journal of Animal Science* 55, 935-942.

Hadjigeorgiou, I., Dardamani, K., Goulas, C., and Zervas, G. (2000). The effect of water availability on feed intake and digestion in sheep. *Small Ruminant Research* 37, 147-150.

Haese, E., Krieg, J., Grubješić, G., Feyder, A., and Rodehutscord, M. (2020). Determination of in situ ruminal degradation of phytate phosphorus from single and compound feeds in dairy cows using chemical analysis and near-infrared spectroscopy. *Animal*, 1-11.

Hafez, E. S. E. (1952). Studies on the breeding season and reproduction of the ewe Part I. The breeding season in different environments Part II. The breeding season in one locality. *The Journal of Agricultural Science* 42, 189-231.

Hagerman, A. D., Thompson, J. M., Ham, C., and Johnson, K. K. (2017). Replacement Beef Cow Valuation under Data Availability Constraints. *Frontiers in Veterinary Science* 4.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, R. L. (2006). Multivariate data analysis 6th Edition. *Pearson Prentice Hall. New Jersey. humans: Critique and reformulation. Journal of Abnormal Psychology* 87, 49-74.

Hamilton, W. J., Sibbald, M. A., and Feist, D. (1995). The effect of sward height on the liveweight gain of farmed yearling red deer (Cervus elaphus) stags. *Grass and Forage Science* 50, 399-404.

Hammack, S. P., and Gill, R. J. (2001). Frame size and weight of cattle. *Bull. L-5176. Texas Agric. Ext. Serv. Texas A & M University*.

Hanks, D. R., Judkins, M. B., McCracken, B. A., Holcombe, D. W., Krysl, L. J., and Park, K. K. (1993). Effects of pregnancy on digesta kinetics and ruminal fermentation in beef cows. *Journal of Animal Science* 71, 2809-2814.

Hatcher, S., Atkins, K. D., and Safari, E. (2009). Phenotypic aspects of lamb survival in Australian Merino sheep1. *Journal of Animal Science* 87, 2781-2790.

Hatcher, S., Eppleston, J., Thornberry, K. J., and Watt, B. (2010). High Merino weaner survival rates are a function of weaning weight and positive post-weaning growth rates. *Animal Production Science* 50, 465-472.

Hausman, J. (2001). Mismeasured variables in econometric analysis: problems from the right and problems from the left. *Journal of Economic perspectives* 15, 57-67.

Hausman, J. A., Newey, W. K., and Powell, J. L. (1995). Nonlinear errors in variables Estimation of some Engel curves. *Journal of Econometrics* 65, 205-233.

Hickson, R. E., Kenyon, P. R., Blair, H. T., Harding, J. E., Oliver, M. H., Jaquiery, A. L., Nicoll, G. B., and Burt, K. G. (2012). The effect of liveweight and liveweight gain of ewes immediately post-weaning on the liveweight and survival of subsequent lambs. *Animal Production Science* 52, 491-496.

Hilario, C. M., Wrage-Mönnig, N., and Isselstein, J. (2017). Behavioral patterns of (co-)grazing cattle and sheep on swards differing in plant diversity. *Applied Animal Behaviour Science* 191, 17-23.

Hinch, G. N., Crosbie, S. F., Kelly, R. W., Owens, J. L., and Davis, G. H. (1985). Influence of birth weight and litter size on lamb survival in high fecundity Booroola-Merino crossbred flocks. *New Zealand Journal of Agricultural Research* 28, 31-38.

Ho, L., Field, R. A., Russell, W. C., Riley, M. L., Ercanbrack, S. K., and Williams Jr, F. L. (1989). Influence of gender, breed and age on maturity characteristics of sheep. *Journal of Animal Science* 67, 2460-2470.

Ho, T. K. (1995). Random decision forests. *Proceedings of 3rd international conference on document analysis and recognition* 1, 278-282.

Hocking, E. J. E., Winslow, E., Behrendt, R., Gordon, D. J., Kearney, G. A., and Thompson, A. N. (2019). Crossbred ewes gain more weight and are fatter than Merino ewes when managed together but similar coefficients predict lamb birthweight and survival. *Animal Production Science* 59, 767-777.

Hocking., E. J. E., Copping, K. J., and Thompson, A. N. (2011). Managing the nutrition of twin-bearing ewes during pregnancy using Lifetimewool recommendations increases production of twin lambs. *Animal Production Science* 51, 813-820.

Hodgson, J., Matthews, P. N. P., Matthews, C., and Lucas, R. J. (1999). Pasture measurement *In* "New Zealand pasture and crop science" (J. White and J. Hodgson, eds.), pp. 56–65. Oxford University Press, Auckland and Melbourne.

Hogan, J. P., Petherick, J. C., and Phillips, C. J. (2007). The physiological and metabolic impacts on sheep and cattle of feed and water deprivation before and during transport. *Nutrition Research Reviews* 20, 17-28.

Holman, B. W. B., Kashani, A., and Malau-Aduli, A. E. O. (2012). Growth and body conformation responses of genetically divergent Australian sheep to Spirulina (Arthrospira platensis) supplementation. *American Journal of Experimental Agriculture* 2, 160-173.

Holst, P. J., Fogarty, N. M., and Stanley, D. F. (2002). Birth weights, meningeal lesions, and survival of diverse genotypes of lambs from Merino and crossbred ewes. *Australian Journal of Agricultural Research* 53, 175-181.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika* 30, 179-185.

Hossamo, H. E., Owen, J. B., and Farid, M. F. A. (1986). Body condition score and production in fat tailed Awassi sheep under range conditions. *Research and Development in Agriculture*.

Hossin, M., and Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process* 5, 1-11.

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical journal* 50, 346-363.

Hughes, G. P., and Harker, K. W. (1950). The technique of weighing bullocks on summer grass. *The Journal of Agricultural Science* 40, 403-409.

Hughes, J. G. (1976). Short-term variation in animal live weight and reduction of its effect on weighing. *Animal Breeding Abstracts* 44, 11.

Imandoust, S. B., and Bolandraftar, M. (2013). Application of k-nearest neighbor (knn) approach for predicting economic events: Theoretical background. *International Journal of Engineering Research and Applications* 3, 605-610.

Jainudeen, M., Wahid, H., and Hafez, E. (2000). Sheep and Goats. *In* "Reproduction in Farm Animals" (B. Hafez and E. Hafez, eds.), pp. 172-181.

Jefferies, B. (1961). Body condition scoring and its use in management. *Tasmanian Journal of Agriculture* 32, 19-21.

Jones, A., van Burgel, A. J., Behrendt, R., Curnow, M., Gordon, D. J., Oldham, C. M., Rose, I. J., and Thompson, A. N. (2011). Evaluation of the impact of Lifetimewool on sheep producers. *Animal Production Science* 51, 857-865.

Junkuszew, A., and Ringdorfer, F. (2005). Computer tomography and ultrasound measurement as methods for the prediction of the body composition of lambs. *Small Ruminant Research* 56, 121-125.

Kaske, M., and Groth, A. (1997). Changes in factors affecting the rate of digesta passage during pregnancy and lactation in sheep fed on hay. *Reproduction Nutrition Development* 37, 573-588.

Keady, T., Mayne, C., Kilpatrick, D., and McCoy, M. (2005). Effect of level and source of nutrients in late gestation on subsequent milk yield and composition and fertility of dairy cows. *Livestock Production Science* 94, 237-248.

Kelly, R. W., and Johnstone, P. D. (1982). Reproductive performance of commercial sheep flocks in South Island districts. *New Zealand Journal of Agricultural Research* 25, 519-523.

Kelly, R. W., Macleod, I., Hynd, P., and Greeff, J. (1996). Nutrition during fetal life alters annual wool production and quality in young Merino sheep. *Australian Journal of Experimental Agriculture* 36, 259-267.

Kennedy, P. M. (1985). Influences of cold exposure on digestion of organic matter, rates of passage of digesta in the gastrointestinal tract, and feeding and rumination behaviour in sheep given four forage diets in the chopped, or ground and pelleted form. *British Journal of Nutrition* 53, 159-173.

Kennedy, P. M., Christopherson, R. J., and Milligan, L. P. (1976). The effect of cold exposure of sheep on digestion, rumen turnover time and efficiency of microbial synthesis. *British Journal of Nutrition* 36, 231-242.

Kenyon, P. R., Hickson, R. E., Hutton, P. G., Morris, S. T., Stafford, K. J., and West, D. M. (2012a). Effect of twin-bearing ewe body condition score and late pregnancy nutrition on lamb performance. *Animal Production Science* 52, 483-490.

Kenyon, P. R., Maloney, S. K., and Blache, D. (2014). Review of sheep body condition score in relation to production characteristics. *New Zealand Journal of Agricultural Research* 57, 38-64.

Kenyon, P. R., Morel, P. C. H., and Morris, S. T. (2004a). Effect of liveweight and condition score of ewes at mating, and shearing mid-pregnancy, on birthweights and growth rates of twin lambs to weaning. *New Zealand Veterinary Journal* 52, 145-149.

Kenyon, P. R., Morel, P. C. H., and Morris, S. T. (2004b). The effect of individual liveweight and condition scores of ewes at mating on reproductive and scanning performance. *New Zealand Veterinary Journal* 52, 230-235.

Kenyon, P. R., Morel, P. C. H., Morris, S. T., Burnham, D. L., and West, D. M. (2006). The effect of length of use of teaser rams prior to mating and individual liveweight on the reproductive performance of ewe hoggets. *New Zealand Veterinary Journal* 54, 91-95.

Kenyon, P. R., Morel, P. C. H., Morris, S. T., and West, D. M. (2005). The effect of individual liveweight and use of teaser rams prior to mating on the reproductive performance of ewe hoggets. *New Zealand Veterinary Journal* 53, 340-343.

Kenyon, P. R., Morris, S. T., Burnham, D. L., and West, D. M. (2008). Effect of nutrition during pregnancy on hogget pregnancy outcome and birthweight and liveweight of lambs. *New Zealand Journal of Agricultural Research* 51, 77-83.

Kenyon, P. R., Morris, S. T., Hickson, R. E., Stafford, K. J., and West, D. M. (2012b). Nutritional restriction of triplet-bearing ewes and body condition score has minimal impacts. *New Zealand Journal of Agricultural Research* 55, 359-370.

Kenyon, P. R., Morris, S. T., Stafford, K. J., and West, D. M. (2011a). Effect of ewe body condition and nutrition in late pregnancy on the performance of triplet-bearing ewes and their progeny. *Animal Production Science* 51, 557-564.

Kenyon, P. R., Morris, S. T., and West, D. M. (2010). Proportion of rams and the condition of ewe lambs at joining influences their breeding performance. *Animal Production Science* 50, 454-459.

Kenyon, P. R., Pain, S. J., Hutton, P. G., Jenkinson, C. M. C., Morris, S. T., Peterson, S. W., and Blair, H. T. (2011b). Effects of twin-bearing ewe nutritional treatments on ewe and lamb performance to weaning. *Animal Production Science* 51, 406-415.

Kenyon, P. R., Stafford, K. J., Jenkinson, C. M. C., Morris, S. T., and West, D. M. (2007a). The body composition and metabolic status of twin- and triplet-bearing ewes and their fetuses in late pregnancy. *Livestock Science* 107, 103-112.

Kenyon, P. R., Stafford, K. J., Jenkinson, C. M. C., Morris, S. T., and West, D. M. (2007b). The body composition and metabolic status of twin-and triplet-bearing ewes and their fetuses in late pregnancy. *Livestock Science* 107, 103-112.

Khaledian, Y., and Miller, B. A. (2020). Selecting appropriate machine learning methods for digital soil mapping. *Applied Mathematical Modelling* 81, 401-418.

Khalifa, E. I., Ahmed, M. E., Hafez, Y. H., El-Zolaky, O. A., Bahera, K. M., and Abido, A. A. (2013). Age at puberty and fertility of Rahmani sheep fed on biological inoculated corn silage. *Annals of Agricultural Sciences* 58, 163-172.

Kim, T. K. (2015). T test as a parametric statistic. *Korean journal of anesthesiology* 68, 540-546.

Kingenberg, C. P. (2003). Growth of Farm Animals. (T. L. J. Lawrence and V. R. Fowler, eds.), pp. 368. CABI Publishing, Wallingford, Oxon.

Kirby, D. R., and Stuth, J. W. (1982). Seasonal diurnal variation in composition of cow diets. *Journal of Range Management* 35, 7-8.

Kirton, A., Moss, R., and Taylor, A. G. (1971). Weight losses from milk and weaned lambs in mid Canterbury resulting from different lengths of starvation before slaughter. *New Zealand Journal of Agricultural Research* 14, 149-160.

Kirton, A. H., Quartermain, A. R., Uljee, A. E., Carter, W. A., and Pickering, F. S. (1968). Effect of 1 and 2 days' ante-mortem fasting on live weight and carcass losses in lambs. *New Zealand Journal of Agricultural Research* 11, 891-902.

Kischel, S. G., Dønnem, I., and Bøe, K. E. (2017). Variation in free water intake in lactating ewes. *Acta Agriculturae Scandinavica, Section A — Animal Science* 67, 117-122.

Kleemann, D. O., and Walker, S. K. (2005). Fertility in South Australian commercial Merino flocks: relationships between reproductive traits and environmental cues. *Theriogenology* 63, 2416-2433.

Knight, T. W. (1980). Effects of diet and liveweight on ovulation rates in Romney ewes. *Proceedings of the New Zealand Society of Animal Production* 40, 38-42.

Kodaira, M., and Shibusawa, S. (2013). Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping. *Geoderma* 199, 64-79.

Koycu, E., Sezenler, T., Ozder, M., and Karadag, O. (2008). The Relationship Between Body Weight and Body Condition Score in Karacabey Merino Ewes. *Journal of Tekiridag Agricultural Faculty* 5, 61–65.

Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software* 28, 1-26.

Lalic, B., Eitzinger, J., Dalla Marta, A., Orlandini, S., Sremac, A. F., and Pacher, B. (2018). "Agricultural Meteorology and Climatology," Firenze University Press, Firenze, Italy.

Lan, Y., Zhou, D., Zhang, H., and Lai, S. (2017). Development of Early Warning Models. *In* "Early Warning for Infectious Disease Outbreak" (W. Yang, ed.), pp. 35-74. Academic Press.

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Lazar, N. (2010). Ockham's razor. *Wiley Interdisciplinary Reviews: Computational Statistics* 2, 243-246.

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *Journal of statistical software* 25, 1-18.

Leach, T. M., and Roberts, C. J. (1981). Present status of chemotherapy and chemoprophylaxis of animal trypanosomiasis in the eastern hemisphere. *Pharmacology & therapeutics* 13, 91-147.

Lechner-Doll, M., Kaske, M., and Engelhardt, W. V. (1991). Factors Affecting the Mean Retention Time of Particles in the Forestomach of Ruminants and Camelids. *In* "Physiological Aspects of Digestion and Metabolism in Ruminants" (T. Tsuda, Y. Sasaki and R. Kawashima, eds.), pp. 455-482. Academic Press, San Diego.

Lechner, I., Barboza, P., Collins, W., Günther, D., Hattendorf, B., Hummel, J., and Clauss, M. (2009). No 'bypass' in adult ruminants: Passage of fluid ingested vs. fluid inserted into the rumen in fistulated muskoxen (Ovibos moschatus), reindeer (Rangifer tarandus) and moose (Alces alces). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* 154, 151-156.

Lee, G. L., Sladek, M. A., Atkins, K. D., and Semple, S. J. (2008). Varience components of liveweights of pregnant ewes measured by manual or remote methods, with and without processing by data screening. *In* "Proceedings of the first joint conference of the Australian and New Zealand Societies of Animal Production (ANZSAP), the Australian Society of Animal Production (ASAP) and the New Zealand Society of Animal Production (NZSAP)" (D. G. Barber, A. R. Cowan and A. R. Anstis, eds.), Vol. 27, pp. 22. Australian Society of Animal Production, University of Queensland: St Lucia, Queensland.

Leevy, J. L., Khoshgoftaar, T. M., Bauder, R. A., and Seliya, N. (2018). A survey on addressing high-class imbalance in big data. *Journal of Big Data* 5, 42.

Li, J. (2017). Assessing the accuracy of predictive models for numerical data: Not r nor r2, why not? Then what? *PloS one* 12, e0183250-e0183250.

Liao, T. F. (1994). Interpreting probability models: Logit, probit, and other generalized linear models. *In* "Quantitative Applications in the Social Sciences", Vol. 101, pp. 96. SAGE Publications.

Lin, L., Dickhoefer, U., Müller, K., Wurina., and Susenbeth, A. (2011). Grazing behavior of sheep at different stocking rates in the Inner Mongolian steppe, China. *Applied Animal Behaviour Science* 129, 36-42.

Litherland, A., Woodward, S., McDougal, D. B., Boom, C., Stevens, D., Knight, T., and Lambert, G. (2002). Seasonal variations in pasture quality in New Zealand sheep and beef farms. *Proceedings of the New Zealand Society of Animal Production* 62, 138-142.

Litherland, A. J., Lambert, M. G., and McLaren, P. N. (1999). Effects of herbage mass and ewe condition score at lambing on lamb survival and liveweight gain. *Proceedings-New Zealand Society of Animal Production* 59, 104-107.

Little, W., Collis, K. A., Gleed, P. T., Sansom, B. F., Allen, W. M., and Quick, A. J. (1980). Effect of reduced water intake by lactating dairy cows on behaviour, milk yield and blood composition. *Vet Rec* 106, 547-51.

Liu, W., Jamshidian, M., and Zhang, Y. (2004). Multiple Comparison of Several Linear Regression Models. *Journal of the American Statistical Association* 99, 395-403.

Lloyd, L. E., McDonald, B. E., and Crampton, E. W. (1978). Fundamentals of nutrition. pp. 466. WH Freeman and Co.

Loureiro, M. F. P., Pain, S., Kenyon, P., Peterson, S., and Blair, H. (2012). Single female offspring born to primiparous ewe-lambs are lighter than those born to adult multiparous ewes but their reproduction and milk production are unaffected. *Animal Production Science* 52, 552-556.

Loureiro, M. F. P., Pain, S. J., Kenyon, P. R., and Blair, H. T. (2010). Do fetuses from primiparous one-year-old ewes differ from those of multiparous mature ewes? *Proceeding of the New Zealand Society of Animal Production.* 70, 118-120.

Martins, B., Mendes, A., Silva, L., Moreira, T., Costa, J., Rotta, P., Chizzotti, M., and Marcondes, M. (2020). Estimating body weight, body condition score, and type traits in dairy cows using three dimensional cameras and manual body measurements. *Livestock Science* 236.

Mathias-Davis, H. C., Shackell, G. H., Greer, G. J., Bryant, A. I., and Everett-Hincks, J. M. (2013). Ewe body condition score and the effect on lamb growth rate. *Proceedings of the New Zealand Society of Animal Production* 73, 131-135.

Maurya, V. P., Sejian, V., Kumar, D., Joshi, A., Naqvi, S. M. K., and Karim, S. A. (2008). Body condition scoring system: a simple tool for optimizing productivity in sheep farms. *CSWRI, Avikanagar*, 1-35.

McDonald, P. (2002). "Animal nutrition," Pearson education.

McDowell, M. L., Bruland, G. L., Deenik, J. L., and Grunwald, S. (2012). Effects of subsetting by carbon content, soil order, and spectral classification on prediction of soil total carbon with diffuse reflectance spectroscopy. *Applied and Environmental Soil Science* 2012.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica* 22, 276-282.

McHugh, N., McGovern, F., Creighton, P., Pabiou, T., McDermott, K., Wall, E., and Berry, D. P. (2019). Mean difference in live-weight per incremental difference in body condition score estimated in multiple sheep breeds and crossbreds. *Animal* 13, 549-553.

McInnes, P., and Smith, M. D. (1966). The effect of nutrition before mating on the reproductive performance of Merino ewes. *Australian Journal of Experimental Agriculture* 6, 455-459.

McMillan, W. H., and Moore, R. W. (1983). Capitalizing on hogget oestrus. *In* "Proceedings of the Society of Sheep and Beef Cattle Veterinarians of the New Zealand Veterinary Association, annual seminar", Vol. (1983).

McNeilly, A. S., Jonassen, J. A., and Rhind, S. M. (1987). Reduced ovarian follicular development as a consequence of low body condition in ewes. *Acta Endocrinologica* 115, 75-83.

Meyer, H. H., and French, R. L. (1979). Hogget liveweight-oestrus relationship among sheep breeds. *Proceedings of the New Zealand Society of Animal Production* 39.

Meyer, J., Lofgreen, G., and Gareett, W. (1960). A proposed method for removing sources of error in beef cattle feeding experiments. *Journal of Animal Science* 19, 1123-1131.

Miaron, J. O. O., and Christopherson, R. J. (1992). Effect of prolonged thermal exposure on heat production, reticular motility, rumen-fluid and -particulate passage-rate constants, and apparent digestibility in steers. *Canadian Journal of Animal Science* 72, 809-819.

Michels, H., Decuypere, E., and Onagbesan, O. (2000). Litter size, ovulation rate and prenatal survival in relation to ewe body weight: genetics review. *Small Ruminant Research* 38, 199-209.

Ministry of Primary Indistries, M. (2018). Code of Welfare: Sheep and Beef Cattle. (N. A. W. A. Committee, ed.), pp. 58, Wellington.

Mir, S., and Ahmed, H. (2017). Effect of Season on the Chemical Composition and Nutritive Value of Shrub Foliage (Cotoneaster spp.) of Sub-Alpine Pasture of Kashmir Valley. *International Journal of Livestock Research* 7, 289-294.

Molina, A., Gallego, L., Perez, J. I., and Bernabeu, R. (1991). Growth of Manchega lambs in relation to body condition of dam, season of birth, type of birth and sex. *Avances en Alimentacion y Mejora Animal*, 31, 198-205.

Morel, P. C. H., Morris, S. T., and Kenyon, P. R. (2009). Effects of birth weight on survival in twin born lambs. *Proceedings of the New Zealand Society of Animal Production* 69, 75-79.

Morel, P. C. H., Schreurs, N. M., Corner-Thomas, R. A., Greer, A. W., Jenkinson, C. M. C., Ridler, A. L., and Kenyon, P. R. (2016). Live weight and body composition associated with an increase in body condition score of mature ewes and the relationship to dietary energy requirements. *Small Ruminant Research* 143, 8-14.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agricultural and Biological Engineers, ASABE* 50, 885-900.

Morley, F. H. W., White, D. H., Kenney, P. A., and Davis, I. F. (1978). Predicting ovulation rate from liveweight in ewes. *Agricultural Systems* 3, 27-45.

Morota, G., Ventura, R. V., Silva, F. F., Koyama, M., and Fernando, S. C. (2018). Big data analytics and precision animal agriculture symposium: Machine learning and data mining advance predictive big data analysis in precision animal agriculture. *Big data analysis in Animal Science* 96, 1540-1550.

Morris, S., and Kenyon, P. (2004). The effect of litter size and sward height on ewe and lamb performance. *New Zealand journal of agricultural research* 47, 275-286.

Morris, S., Parker, W., Burnham, D., and Jenkinson, C. (1993). Herbage allowance-intake-production relationships in cotinuously stocked winter and spring-lambing ewes. *Proceedings-New Zealand Society of Animal Production* 53, 11-11.

Morris, S. T., Kenyon, P. R., and Burnham, D. L. (2002). A comparison of two scales of body condition scoring in Hereford x Friesan beef breeding cows. *Proceedings of the New Zealand Grassland Association* 64, 121-123.

Moyo, M., and Nsahlai, I. (2018). Rate of Passsage of Digesta in Ruminants; Are Goats Different? *In* "Goat Science" (S. Kukovics, ed.), Vol. Open access peer-reviewed Edited Volume, pp. 39-74. IntechOpen.

Naeger, D. M., Kohi, M. P., Webb, E. M., Phelps, A., Ordovas, K. G., and Newman, T. B. (2013). Correctly using sensitivity, specificity, and predictive values in clinical practice: how to avoid three common pitfalls. *American Journal of Roentgenology* 200, W566-W570.

Nawar, S., and Mouazen, A. M. (2017). Predictive performance of mobile vis-near infrared spectroscopy for key soil properties at different geographical scales by using spiking and data mining techniques. *Catena* 151, 118-129.

Newton, J. E., Betts, J. E., and Wilde, R. (1980). The effect of body condition and time of mating on the reproductive performance of Masham ewes. *Animal Science* 30, 253-260.

Nicol, A., and Brookes, I. (2007). The metabolisable energy requirements of grazing livestock. *In* "Pasture and supplements for grazing animals", Vol. 14, pp. 151-172.

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education* 15, 625-632.

Obuchowski, N. A., and Bullen, J. A. (2018). Receiver operating characteristic (ROC) curves: review of methods with applications in diagnostic medicine. *Physics in Medicine & Biology* 63, 07TR01.

Oldham, C. M., Thompson, A. N., Ferguson, M. B., Gordon, D. J., Kearney, G. A., and Paganoni, B. L. (2011). The birthweight and survival of Merino lambs can be predicted from the profile of liveweight change of their mothers during pregnancy. *Animal Production Science* 51, 776-783.

Orr, R. J., Penning, P. D., Harvey, A., and Champion, R. A. (1997). Diurnal patterns of intake rate by sheep grazing monocultures of ryegrass or white clover. *Applied Animal Behaviour Science* 52, 65-77.

Oshita, T., Sudo, K., Nonaka, K., Kume, S., and Ochiai, K. (2008). The effect of feed regimen on chewing time, digesta passage rate and particle size distribution in

Holstein non-lactating cows fed pasture ad libitum. *Livestock Science* 113, 243-250.

Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C., and Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian journal of ophthalmology* 56, 45.

Parra, R. (1978). Comparison of foregut and hindgut fermentation in herbivores. *In* "The Ecology of Arboreal Folivores." (G. Montgomery, ed.), pp. 209-229. Smithsonian Institute Press, Washington DC, USA: .

Peart, J. N. (1970). The influence of live weight and body condition on the subsequent milk production of Blackface ewes following a period of undernourishment in early lactation. *The Journal of Agricultural Science* 75, 459-469.

Peart, J. N. (1982). Lactation of suckling ewes and does. *Sheep and Goat Production. World Animal Science C* 1, 119-134.

Peat, J., and Barton, B. (2008). "Medical Statistics: Aguide to data analysis and critical appraisal," Wiley.

Penning, P., and Hooper, G. (1985). An evaluation of the use of short-term weight changes in grazing sheep for estimating herbage intake. *Grass and Forage Science* 40, 79-84.

Pettigrew, E., Hicksona, R., Blaira, H., Griffithsa, K., Ridlera, A., Morrisa, S., and Kenyona, P. (2018). Differences in birth weight and neonatal survival rate of lambs born to ewe hoggets or mature ewes. *New Zealand Journal of Animal Science and Production* 78 16-20.

Pettigrew, E. J., Hickson, R. E., Morris, S. T., Lopez-Villalobos, N., Pain, S. J., Kenyon, P. R., and Blair, H. T. (2019). The effects of birth rank (single or twin) and dam age on the lifetime productive performance of female dual purpose sheep (Ovis aries) offspring in New Zealand. *PloS one* 14.

Phythian, C. J., Hughes, D., Michalopoulou, E., Cripps, P. J., and Duncan, J. S. (2012). Reliability of body condition scoring of sheep for cross-farm assessments. *Small Ruminant Research* 104, 156-162.

Piggot, G. J. (2009). Low dry matter content of northern dairy pastures. *Proceedings of the New Zealand Grassland Association* 71, 31-34.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and Team., R. C. (2018). nlme: Linear and Nonlinear Mixed Effects Models. *In* "R package version 3.1-137". Available at https://cran.r-project.org/web/packages/nlme/nlme.pdf. Accessed on 4/23/2018.

Pischke, S. (2007). Lecture notes on measurement error. London School of Economics, London. available at https://economics.lse.ac.uk/staff/spischke/ec524/Merr_new.pdf, Accessed on 16/11/2018.

Pollott, G. E., and Kilkenny, J. B. (1976). A note on the use of condition scoring in commercial sheep flocks. *Animal Science* 23, 261-264.

Poole, M. A., and O'Farrell, P. N. (1971). The Assumptions of the Linear Regression Model. *Transactions of the Institute of British Geographers*, 145-158.

Quinlan, J. R. (1986). Induction of decision trees. *Machine learning* 1, 81-106.

Quinlan, J. R. (1987). Simplifying decision trees. *International Journal of Man-Machine Studies* 27, 221-234.

Quirke, F. J., Stabenfeldt, G. H., and Bradford, G. E. (1985). "Onset of Puberty and Duration of the Breeding Season in Suffolk, Rambouillet, Finnish Landrace, Dorset and Finn-Dorset Ewe Lambs," pp. 1-1463-71.

R Core Team (2016). R: A language and environment for statistical computing. . R Foundation for Statistical Computing. R version 3.4.4 (2018-03-15), Vienna, Austria. Available at https://cran.r-project.org/.

Rahe, A. J. (1974). Tables of critical values for the Pratt matched pair signed rank statistic. *Journal of the American Statistical Association* 69, 368-373.

Rattray, P. V., Garrett, W. N., East, N. E., and Hinman, N. (1974). Growth, development and composition of the ovine conceptus and mammary gland during pregnancy. *Journal of Animal Science* 38, 613-626.

Rennie, J. D., Shih, L., Teevan, J., and Karger, D. R. (2003). Tackling the poor assumptions of naive bayes text classifiers. *In* "Proceedings of the 20th international conference on machine learning (ICML-03)", pp. 616-623.

Rhind, S. M., Doney, J. M., Gunn, R. G., and Leslie, I. D. (1984a). Effects of body condition and environmental stress on ovulation rate, embryo survival, and associated plasma follicle stimulating hormone, luteinizing hormone, prolactin and progesterone profiles in Scottish Blackface ewes. *Animal Science* 38, 201-209.

Rhind, S. M., Gunn, R. G., Doney, J. M., and Leslie, I. D. (1984b). A note on the reproductive performance of greyface ewes in moderately fat and very fat condition at mating. *Animal Science* 38, 305-307.

Richards, J. S., Atkins, K. D., Thompson, T., and Murray, W. K. (2006). Data from walk-over weighing–where are we at? *In* "Proceedings from the Australian Society of Animal Production 26th Biennial Conference. Short Communication 32" (N. R. Adams, K. P. Croker and D. R. Lindsay, eds.). The Australian Society of Animal Production, University of Western Australia, Perth.

Ridler, A. L., Corner-Thomas, R. A., Kenyon, P. R., and Griffiths, K. J. (2017). Investigation of fetal loss in ewe lambs in relation to liveweight changes and progesterone concentrations in early to mid gestation. *New Zealand Veterinary Journal* 65, 34-38.

Roche, J., Berry, D., and Kolver, E. (2006). Holstein-Friesian strain and feed effects on milk production, body weight, and body condition score profiles in grazing dairy cows. *Journal of dairy science* 89, 3532-3543.

Rondon, Z., Forcada, F., Zarazaga, L., Abecia, J. A., and Lozano, J. M. (1996). Oestrous activity, ovulation rate and plasma melatonin concentrations in Rasa Aragonesa ewes maintained at two different and constant body condition score levels and implanted or reimplanted with melatonin. *Animal Reproduction Science* 41, 225-236.

Rook, A. J., and Penning, P. D. (1991). Synchronisation of eating, ruminating and idling activity by grazing sheep. *Applied Animal Behaviour Science* 32, 157-166.

Rosales, N. C. A., Ferguson, M. B., Macleay, C. A., Briegel, J. R., Martin, G. B., and Thompson, A. N. (2013). Selection for superior growth advances the onset of puberty and increases reproductive performance in ewe lambs. *Animal* 7, 990-997.

Roughan, P. G., and Holland, R. (1977). Predicting in-vivo digestibilities of herbages by exhaustive enzymic hydrolysis of cell walls. *Journal of the Science of Food and Agriculture* 28, 1057-1064.

Rowe, J. B. (2003). Nutritional management of the Australian sheep flock. *Recent Advances in Animal Nutrition in Australia* 14, 23-31.

Rozeboom, K. J., Neale, B. D., and Darroch, C. S. (2007). Relationships among ewe body condition scores, lamb vigor, colostral quality, milk composition and reproductive performance. *Journal of Animal Science* 85, 138-138.

Rueda, B. R., Hallford, D. M., Holcombe, D. W., Broesder, J. T., Krysl, L. J., Judkins, M. B., Barton, R. K., and Gunter, S. A. (1990). Digesta kinetics, ruminal fermentation characteristics and serum metabolites of pregnant and lactating ewes fed chopped alfalfa hay. *Journal of Animal Science* 68, 3821-3831.

Russel, A. (1984). Body condition scoring of sheep. *In Practice* 6, 91-93.

Russel, A. J. F., Doney, J. M., and Gunn, R. G. (1969). Subjective assessment of body fat in live sheep. *The Journal of Agricultural Science* 72, 451-454.

Russell, L. (2018). Emmeans: estimated marginal means, aka least-squares means. *In* "R package version 1.5.4". Available at https://github.com/rvlenth/emmeans.

Rutherford, L., Nicol, A. M., and Logan, C. M. (2003). Recognising the limits to live weight reproduction relationships in ewes. *In* "Proceedings-New Zealand Society of Animal Production", Vol. 63, pp. 140-143. New Zealand Society of Animal Production; 1999.

Sanson, D. W., West, T. R., Tatman, W. R., Riley, M. L., Judkins, M. B., and Moss, G. E. (1993). Relationship of body composition of mature ewes with condition score and body weight. *Journal of Animal Science* 71, 1112-1116.

Särkijärvi, S., Niemeläinen, O., Sormunen-Cristian, R., and Saastamoinen, M. (2012). Changes in chemical composition of different grass species and -mixtures in equine pasture during grazing season. *In* "Forages and grazing in horse nutrition" (M. Saastamoinen, M. J. Fradinho, A. S. Santos and N. Miraglia, eds.), pp. 45-48. Wageningen Academic Publishers, Wageningen.

Saul, G., Kearney, G., and Borg, D. (2011). Pasture systems to improve productivity of sheep in south-western Victoria 2. Animal production from ewes and lambs. *Animal Production Science* 51, 982-989.

Scaramuzzi, R. J., Campbell, B. K., Downing, J. A., Kendall, N. R., Khalid, M., Muñoz-Gutiérrez, M., and Somchit, A. (2006). A review of the effects of supplementary nutrition in the ewe on the concentrations of reproductive and metabolic hormones and the mechanisms that regulate folliculogenesis and ovulation rate. *Reproduction Nutrition Development* 46, 339-354.

Schreurs, N. M., Kenyon, P. R., Morel, P. C. H., and Morris, S. T. (2012). Meta-analysis to establish the response of having heavier mature ewes during gestation on the birthweight of the lamb and the weaning weight of the ewe and lamb. *Animal Production Science* 52, 540-545.

Sejian, V., Maurya, V. P., Naqvi, S. M. K., Kumar, D., and Joshi, A. (2010). Effect of induced body condition score differences on physiological response, productive and reproductive performance of Malpura ewes kept in a hot, semi-arid environment. *Journal of Animal Physiology and Animal Nutrition* 94, 154-161.

Semakula, J., Corner-Thomas, R., Morris, S., Blair, H., and Kenyon, P. (2019). The effect of herbage type prior to fasting on the rate of liveweight loss during fasting in ewe lambs. *New Zealand Journal of Animal Science and Production* 79, 131-134.

Semakula, J., Corner-Thomas, R. A., Morris, S. T., Blair, H. T., and Kenyon, P. R. (2020a). The Effect of Age, Stage of the Annual Production Cycle and Pregnancy-Rank on the Relationship between Liveweight and Body Condition Score in Extensively Managed Romney Ewes. *Animals* 10, 784.

Sezenler, T., Özder, M., Yildirir, M., Ceyhan, A., and Yüksel, M. A. (2011). The relationship between body weight and body condition score in some indigenous sheep breeds in Turkey. *Journal of Animal and Plant Science* 21, 443-447.

Shahinfar, S., and Kahn, L. (2018). Machine learning approaches for early prediction of adult wool growth and quality in Australian Merino sheep. *Computers and Electronics in Agriculture* 148, 72-81.

Shahinfar, S., Kelman, K., and Kahn, L. (2019). Prediction of sheep carcass traits from early-life records using machine learning. *Computers and Electronics in Agriculture* 156, 159-177.

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591-611.

Shrestha, J. N. B., Heaney, D. P., and Vesely, J. A. (1991). Full-feeding versus overnight fasting for assessing growth potential of lambs. *Small Ruminant Research* 5, 117-127.

Silanikove, N. (1992). Effects of water scarcity and hot environment on appetite and digestion in ruminants: a review. *Livestock Production Science* 30, 175-194.

Silva, S. R., Payan-Carreira, R., Quaresma, M., Guedes, C. M., and Santos, A. S. (2016). Relationships between body condition score and ultrasound skin-associated subcutaneous fat depth in equids. *Acta Veterinaria Scandinavica* 58, 62.

Smith, J. F. (1982). Principles of reproduction. *Sheep Production: Breeding and Reproduction, New Zealand Institute of Agricultural Science*, 211-237.

Smith, R. A., and Turner, M. J. B. (1974). Electronic aids for use in fastock weighing. *Journal of Agricultural Engineering Research* 19, 299-311.

Smith, R. N. (1956). Fusion of the epiphyses of the limb bones of the sheep. *The Veterinary Record*. 68, 257-259.

Somasiri, S. C., Kenyon, P. R., Kemp, P. D., Morel, P. C. H., and Morris, S. T. (2016). Effect of herb-clover mixes of plantain and chicory on yearling lamb production in the early spring period. *Animal Production Science* 56, 1662-1668.

Stephenson, S. K., Dalton, D. C., and Kirton, A. H. (1980). The relationship of growth; body shape, and body composition to the initiation of oestrous activity in different sheep breeds. *Proceedings of the New Zealand Society of Animal Production* 40, 258-267.

Story, L. F., and Ross, D. A. (1960). Effect of shearing time on wool. *New Zealand Journal of Agricultural Research* 3, 113-124.

Suiter, J. (1994). Body condition scoring of sheep and goats. *Farmnote* 69, 1994.

Sullivan, G. M., and Artino, A. R. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of graduate medical education* 5, 541-542.

Sumner, R. M. W., Clarke, J. N., Pearson, A. J., and Speedy, P. M. (1994). Repeatability of seasonal wool growth in Merino Sheep. *In* "Proceedings of the New Zealand Society of Animal Production", Vol. 54, pp. 163-166. New Zealand Society of Animal Production.

Sun, S., and Huang, R. (2010). An adaptive k-nearest neighbor algorithm. *In* "2010 seventh international conference on fuzzy systems and knowledge discovery", Vol. 1, pp. 91-94. Institute of Electrical and Electronics Engineers. IEEE.

Taylor, B. N., and Kuyatt, C. E. (1994). "Guidelines for evaluating and expressing the uncertainty of NIST measurement results. NIST technical note: 1297." Department of Commerce, Technology Administration, National Institute of Standards and Technology, Gaithersburg, MD, U.S. .

Teixeira, A., Delfa, R., and Colomer-Rocher, F. (1989). Relationships between fat depots and body condition score or tail fatness in the Rasa Aragonesa breed. *Animal Science* 49, 275-280.

Tharwat, A. (2020). Classification assessment methods. *New England Journal of Entrepreneurship* ahead-of-print. Available at https://www.emerald.com. Accessed on 10/12/2020.

Theil, H. (1958). Economic policy and forecasts. assisted by J.S. Cramer. H. Moerman and A. Russchen. *In* "Contributions to economic analysis", pp. 562. North-Holland Pub.Co, Amsterdam.

Thompson, A. N., Ferguson, M. B., Campbell, A. J. D., Gordon, D. J., Kearney, G. A., Oldham, C. M., and Paganoni, B. L. (2011). Improving the nutrition of Merino ewes during pregnancy and lactation increases weaning weight and survival of progeny but does not affect their mature size. *Animal Production Science* 51, 784-793.

Thompson, A. N., and Oldham, C. M. (2004). Lifetime wool. 5. Carryover effects on ewe reproduction. *Science Access* 1, 291-291.

Thompson, J., O'Halloran, W., McNeill, D., Jackson-Hope, N., and May, T. (1987). The effect of fasting on liveweight and carcass characteristics in lambs. *Meat Science* 20, 293-309.

Thompson, J. M., and Meyer, H. H. (1994). Body condition scoring of sheep. *Proceedings of Western Section American Society of Animal Science* 43, 175.

Thomson, B. C., Muir, P. D., and Smith, N. B. (2004). Litter size, lamb survival, birth and twelve week weight in lambs born to cross-bred ewes. *Proceedings of the New Zealand Grassland Association* 66, 233-237.

Tibshirani, R. J., and Tibshirani, R. (2009). A bias correction for the minimum error rate in cross-validation. *The Annals of Applied Statistics*, 822-829.

Torgo, L. (2016). "Data mining with R: learning with case studies," CRC press.

Toupet, R., Gibbons, A. T., Goodacre, S. L., and Bell, M. J. (2020). Effect of herbage density, height and age on nutrient and invertebrate generalist predator abundance in permanent and temporary pastures. *Land* 9, 164.

Triguero, I., del Río, S., López, V., Bacardit, J., Benítez, J. M., and Herrera, F. (2015). ROSEFW-RF: the winner algorithm for the ECBDL'14 big data competition: an extremely imbalanced big data bioinformatics problem. *Knowledge-Based Systems* 87, 69-79.

Tropsha, A., Gramatica, P., and Gombar, V. K. (2003). The Importance of Being Earnest: Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. *QSAR & Combinatorial Science* 22, 69-77.

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of clinical epidemiology* 49, 1225-1231.

Valletta, J. J., Torney, C., Kings, M., Thornton, A., and Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour* 124, 203-220.

van Burgel, A. J., Oldham, C. M., Behrendt, R., Curnow, M., Gordon, D. J., and Thompson, A. N. (2011). The merit of condition score and fat score as alternatives to liveweight for managing the nutrition of ewes. *Animal Production Science* 51, 834-841.

Van Der Linden, D. S., Kenyon, P. R., Blair, H. T., Lopez-Villalobos, N., Jenkinson, C. M. C., Peterson, S. W., and Mackenzie, D. D. S. (2009). Effects of ewe size and nutrition on fetal mammary gland development and lactational performance of offspring at their first lactation. *Journal of Animal Science* 87, 3944-3954.

van der Linden, D. S., Kenyon, P. R., Blair, H. T., Lopez-Villalobos, N., Jenkinson, C. M. C., Peterson, S. W., and Mackenzie, D. D. S. (2010). Relationships between early postnatal growth and metabolic function of 16-month-old female offspring born to ewes exposed to different environments during pregnancy. *Journal of Developmental Origins of Health and Disease* 1, 50-59.

Van Weyenberg, S., Sales, J., and Janssens, G. P. J. (2006). Passage rate of digesta through the equine gastrointestinal tract: A review. *Livestock Science* 99, 3-12.

Varga, G. A., and Prigge, E. C. (1982). Influence of forage Species and Level of Intake on Ruminal Turnover Rates. *Journal of Animal Science* 55, 1498-1504.

Venables, W. N., and Ripley, B. D. (2002). "Modern applied statistics with S," 4th/Ed. Springer, New York, pp. 1-495.

Verbeek, E., Waas, J. R., Oliver, M. H., McLeay, L. M., Ferguson, D. M., and Matthews, L. R. (2012). Motivation to obtain a food reward of pregnant ewes in negative energy balance: Behavioural, metabolic and endocrine considerations. *Hormones and Behavior* 62, 162-172.

Vieira, A., Brandão, S., Monteiro, A., Ajuda, I., and Stilwell, G. (2015). Development and validation of a visual body condition scoring system for dairy goats with picture-based training. *Journal of Dairy Science* 98, 6597-6608.

Wamatu, J., Alkhtib, A., and Rischkowsky, B. (2019). Simple and robust model to estimate liveweight of Ethiopian Menz sheep. *Animal Production Science* 59, 2265-2272.

Warly, L., Fariani, A., Ichinohe, T., and Fujihara, T. (2004). Seasonal changes in nutritive value of some grass species in West Sumatra, Indonesia. *Asian-Australasian Journal of Animal Sciences* 17, 1663-1668.

Wheeler, J. L., Reardon, T. F., Hedges, D. A., and Rocks, R. L. (1971). The contribution of the conceptus to weight change in pregnant Merino ewes at pasture. *The Journal of Agricultural Science* 76, 347-353.

Whiteman, J. V., Loggins, P. F., Chambers, D., Pope, L. S., and Stephens, D. F. (1954). Some sources of error in weighing steers off grass. *Journal of Animal Science* 13, 832-842.

Wicker, J. E. (2006). Applications of modern statistical methods to analysis of data in physical science. PhD dissertion, University of Tennessee, Knoxville, pp. 1-247.

Wiener, G. (1967). A comparison of the body size, fleece weight and maternal performance of five breeds of sheep kept in one environment. *Animal Science* 9, 177-195.

Wilson, R. (2014). Optimising a weighing protocol for sheep, Unpublished honours dissertation, Lincoln University, pp. 1- 57.

Wilson, R. F., Hardwick-Smith, J. A., Logan, C. M., Corner-Thomas, R. A., Bywater, A. C., and Greer, A. W. (2015). Brief communication: Optimizing a fasting protocol to assess live weight of sheep. *Proceedings of the New Zealand Society of Animal Production* 75, 223-224.

Wishart, H., Morgan-Davies, C., Stott, A., Wilson, R., and Waterhouse, T. (2017). Liveweight loss associated with handling and weighing of grazing sheep. *Small Ruminant Research* 153, 163-170.

Yates, W. J., and Gleeson, A. R. (1975). Relationships between condition score and carcass composition of pregnant Merino sheep. *Australian Journal of Experimental Agriculture* 15, 467-470.

Young, B. A. (1981). Cold Stress as it Affects Animal Production. *Journal of Animal Science* 52, 154-163.

Young, B. A., and Corbett, J. L. (1972). Maintenance energy requirement of grazing sheep in relation to herbage availability. I. Calorimetric estimates. *Australian Journal of Agricultural Research* 23, 57-76.

Young, J., Thompson, A., and Oldham, C. (2004). Lifetime wool. 15. Whole-farm benefits from optimising lifetime wool production. *In* "Proceedings of the 25th Biennial Conference of the Australian Society of Animal Production", Vol. 25, pp. 338, Melbourne. Available at http://www.lifetimewool.com.au/pdf/lw_asap15.pdf.

Yu, H., and Yang, J. (2001). A direct LDA algorithm for high-dimensional data—with application to face recognition. *Pattern recognition* 34, 2067-2070.

Zarkawi, M., and Al-Daker, A.-M. B. (2016). Body weight and reproductive parameters in fast and weak growing Awassi ram lambs during different age stages. *Tropical Animal Health and Production* 48, 223-227.

Zeng, Z.-Q., Yu, H.-B., Xu, H.-R., Xie, Y.-Q., and Gao, J. (2008). Fast training support vector machines using parallel sequential minimal optimization. *In* "2008 3rd international conference on intelligent system and knowledge engineering", Vol. 1, pp. 997-1001. International Journal of Pattern Recognition and Artificial Intelligence, IEEE.

Zhang, L., and Zhan, C. (2017). Machine learning in rock facies classification: an application of XGBoost. *In* "International Geophysical Conference" (Y. Wang, and and Z. Wang, eds.), pp. 1371-1374. Society of Exploration Geophysicists and Chinese Petroleum Society, Qingdao, China.

Zhao, H., Wang, Z., and Nie, F. (2018). A new formulation of linear discriminant analysis for robust dimensionality reduction. *In* "IEEE Transactions on Knowledge and data engineering", Vol. 31, pp. 629-640.

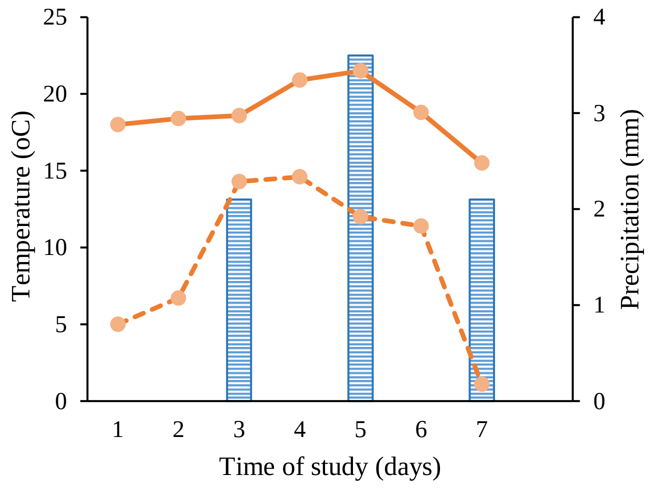
Zheng, W., Zhao, L., and Zou, C. (2004). An efficient algorithm to solve the small sample size problem for LDA. *Pattern Recognition* 37, 1077-1079.

Zhu, F., Tang, M., Xie, L., and Zhu, H. (2018). A Classification Algorithm of CART Decision Tree based on MapReduce Attribute Weights. *International Journal of Performability Engineering* 14.

Zulu, V. C., Nakao, T., Moriyoshi, M., Nakada, K., Sawamukai, Y., Tanaka, Y., and Zhang, W.-C. (2001). Relationship between body condition score and ultrasonographic measurement of subcutaneous fat in dairy cows. *Asian-Australasian Journal of Animal Sciences* 14, 816-820.

Appendices

Appendix I. Weather data for Chapter 3



Appendix I Figure 1 Average daily temperature (solid line: maximum, dashed: minimum) and precipitation (stripped bars) for the study time. Data from: https://cliflo.niwa.co.nz.

Appendix II. Water availability study: Effect of water availability to liveweight loss rate of fasting ewe lambs

Introduction

The lambs in Chapter 3 were part of a larger study that compared the impacts of herbage type (grass and herb) with or without access to a reticulated water trough (corner et al unpublished), although the analysis in Chapter 3 focused only on herbage type. The aim of the present study was to investigate the effect of access to reticulated water on the rate of liveweight loss in lambs when removed from herbage.

Materials and methods

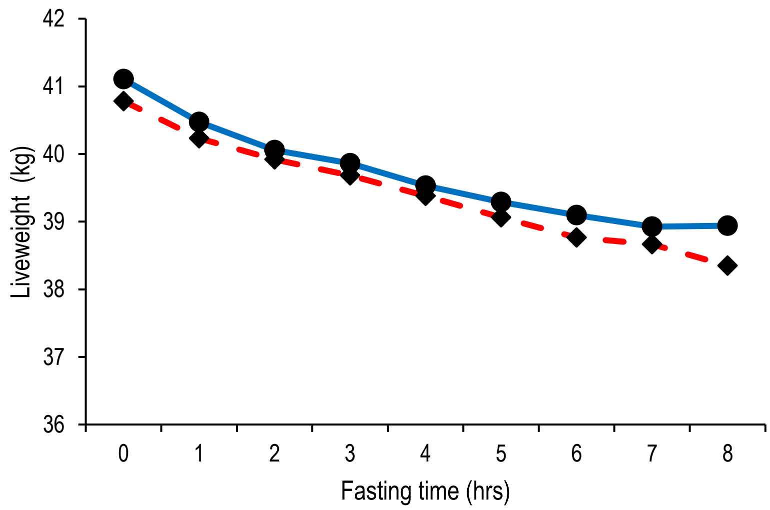
This trial was part of an ongoing study (Corner-Thomas et al, unpublished) conducted simultaneously with the herbage type study reported in Chapter 6 in which additional data on effect of access to reticulated water was collected. In their study, Corner-Thomas et al (unpublished) allocated six-month-old ewe lambs (n = 80) to one of two dietary treatments: an established ryegrass and white clover dominated sward (grass, G) or a chicory, plantain, red and white clover mix (herb-clover, H). They were also further, allocated to one of two water treatments: no access to reticulated water (NW) or access to reticulated water (W). This allocation resulted in four treatment combinations: grass without reticulated water (GNW, n = 20), grass with reticulated water (GW, n = 20), herb-clover without reticulated water (HNW, n = 20) and herb-clover with reticulated water (HW, n = 20). To investigate the effect of water access, data were pooled by water treatment group namely, access to reticulated water (AW, n = 40) and no access (NW = 40), each replicated twice (i.e. by herbage type). The lambs were maintained on these treatments for 30 days prior to weighing. All the experimental and data collection conditions were as reported in Chapter 6.

Statistical analyses

Analyses were conducted using R program version 3.4.4 (R Core Team, 2016). The analytical methodology utilised in this study were similar to ones applied in Chapter 3. A linear mixed-effects model with polynomial time effect was fitted using nlme, a package for fitting regression for linear and nonlinear models (Pinheiro et al., 2018). Effects in the model were contrasted based on successive differences comparison (Liu et al., 2004) using the MASS package (Venables and Ripley, 2002). Access to water was fitted as a fixed variable, fasting time (linear and quadratic) as a covariate while an individual sheep effect was fitted as a random effect. Two-way interactions of access to water x fasting time were also fitted. An autoregressive correlation structure was fitted, to account for temporal dependency of nearby time.

Results and discussion

Initial liveweights of lambs in the AW and NW treatment groups were 40.8±1.10 kg and 41.1±0.82 kg, while their final liveweights were 38.9±0.99 kg and 38.3±0.86 kg, respectively (Appendix II Figure 1). Lambs in AW and NW treatment groups lost significant amounts of liveweight after four (1.4±0.15 kg, and 1.6±0.14 kg or 3.4% and 3.8% liveweight) and eight hours (2.4±0.12 and 2.2±0.12 kg or 5.8% and 5.3% liveweight), respectively. Access to drinking water, had no effect ($p > 0.05$) on the rate of liveweight loss over the entire holding time (Appendix II Table 1, Appendix II Figure 2). Therefore, data were pooled for the two water treatment groups to generate an overall prediction equation.

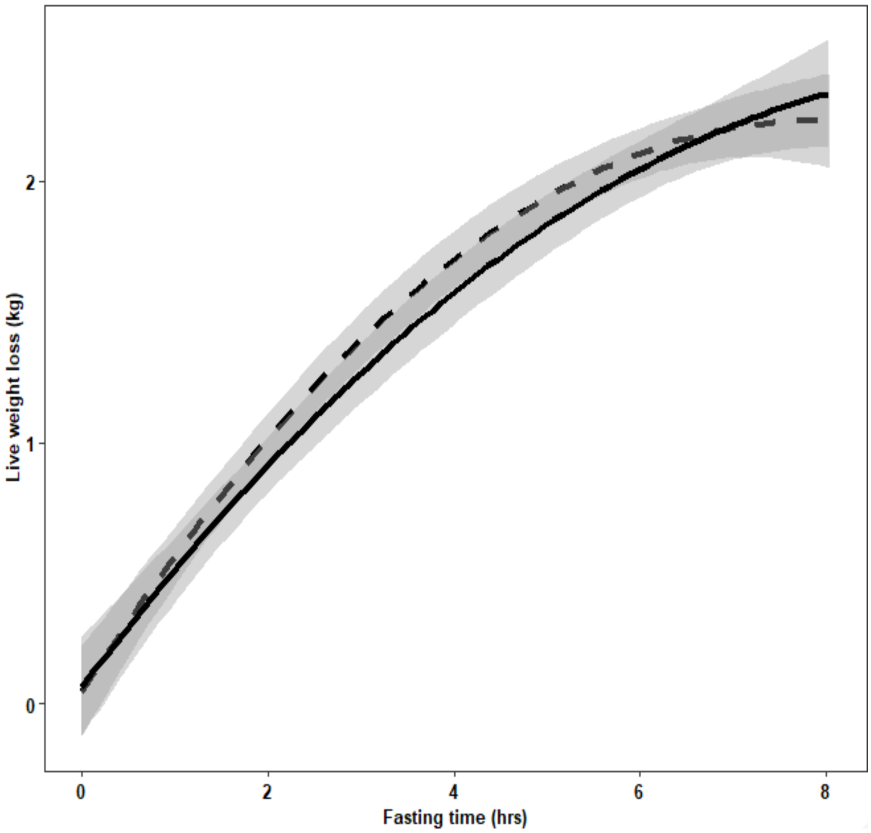


APPENDIX II Figure 1 Plot of liveweight decay for access to water (solid line) and no access to water (dashed line) treatments.

APPENDIX II Table 1 Prediction parameters with standard errors in parentheses for lamb liveweight loss (kg) for the water access treatments (AW and NW).

	Predictor			
Water access	Intercept	Time	Time^2	Adjusted R^2
AW	0.05 (0.098)	0.55 (0.038)	−0.035 (0.005)	0.75
NW	0.07 (0.111)	0.47 (0.037)	−0.020 (0.005)	0.79
Overall[a]	0.06 (0.073)	0.51 (0.026)	−0.030 (0.003)	0.78

[ab] Overall liveweight loss prediction equation ($LWL=0.06+0.51Time - 0.030Time^2$, $R^2 = 0.78$).



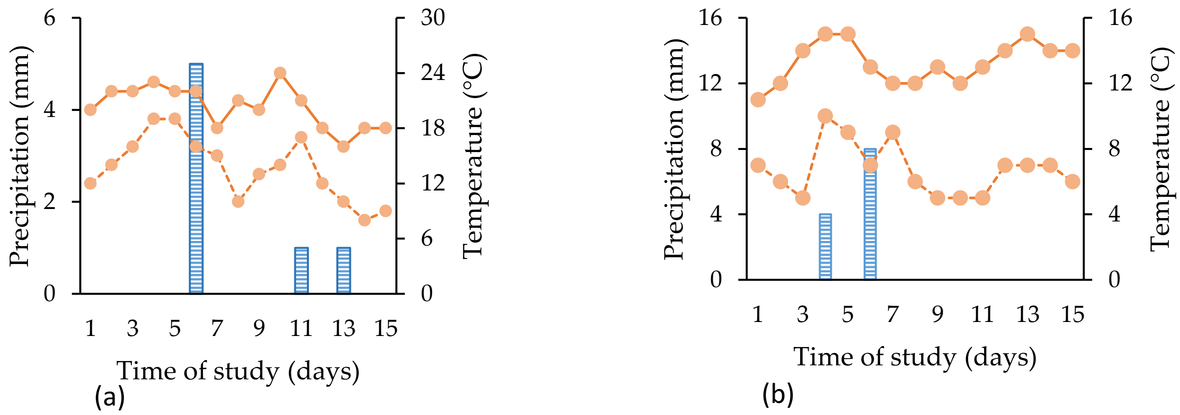
APPENDIX II Figure 2 Change in liveweight (with 95% Confidence Interval, grey shade) after removal from herbage, for water (AW: solid line) and no water (NW: dashed line) treatments.

In the current study, restricted access to reticulated water had no effect on the overall liveweight loss or rate of liveweight loss of the lambs. This finding is in agreement with Kirton et al. (1968) and Al-Ramamneh et al. (2012) who reported no difference in overall liveweight loss regardless of access to drinking water or not among Romney lambs managed under extensive grazing conditions and two temperate sheep breeds kept under zero grazing. These results were also in agreement with studies under different environmental conditions with adult sheep who reported no difference in overall liveweight loss across two water access (access, no access) treatments (Brosh et al., 1986; Hadjigeorgiou et al., 2000). In many studies, liveweight loss in ruminants was associated with a reduction in water and feed intake and was influenced by ambient environmental temperatures and body-water loss (Silanikove, 1992; Alamer, 2009). Given the low dry matter (DM) percentage in the present study (AW: 16.7%, NW: 16.8%) the lambs were likely to have consumed a considerable amount of water in their herbage intake. Lambs need to consume dry matter % of approximately 2 to 3% of their liveweight (Lloyd et al., 1978; McDonald, 2002). In the current study all lambs were offered *ad libitum* pasture levels >1200 kg DM/ha, therefore, the average daily intake of an average lamb (41.0 kg) was estimated at 1.23 (0.03 x 41.0) kg DM/ha. Therefore, if the herbage had a dry matter of 16.7%, lambs would have consumed approximately 6.10 (1.23 kg DM/ha x 4.94) litres of water/ha assuming a water density of 1.0 kg/litre. The study was conducted when the ambient temperatures were relatively low (Average: 13°C, range: 9 to 19°C) which could have reduced their need for drinking water. It was unknown if the lambs allowed access to drinking water, had actually been drinking water which may warrant further research.

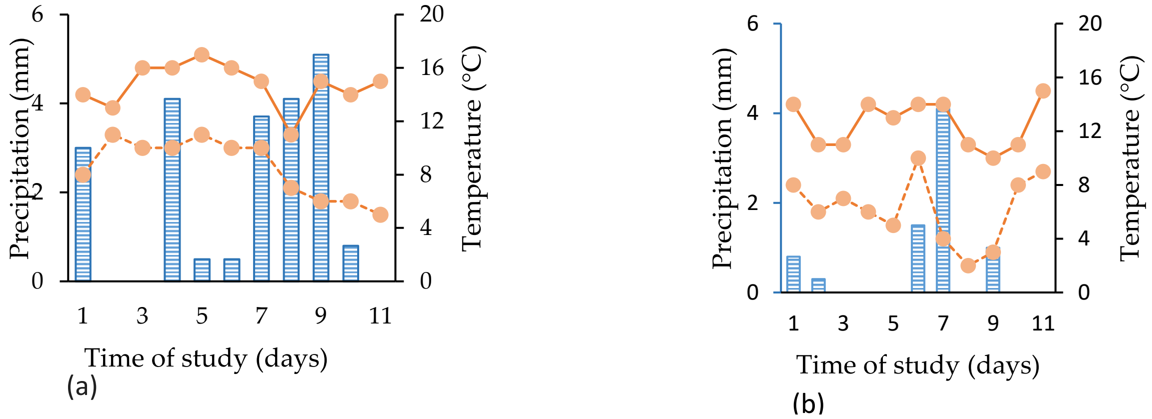
Conclusions

The present study identified that the liveweight loss profile during an eight-hour fast did not differ for lambs given access to reticulated water or not prior to fasting. Therefore, when adjusting lamb liveweight for losses associated with the duration after removal from paddock, whether lambs had access to drinking water or not need not to be considered.

Appendix III. Weather data for Chapter 4.



Appendix III Figure 1 Average daily precipitation (stripped bars) and temperature (solid line: maximum, dashed: minimum) during the calibration stage over the study time in autumn (a) and winter (b).



Appendix III Figure 2 Daily temperature (solid line: maximum, dashed: minimum) and precipitation (stripped bars) during the validation stage over the study time for Tuapaka farm (a) and Riverside farm (b).

Data from: https://cliflo.niwa.co.nz.

Appendix IV. Herbage mass and chemical composition for Chapter 4

Appendix IV Table 1a Estimated post feeding herbage mass (least squares means) and proportion of live dry matter (%) of Low, Medium, and High herbage availability levels (kg DM/ha) offered to ewe by season (autumn, winter) during calibration.

Study	Herbage availability	Herbage mass (kg DM/ha)				Proportion of live dry matter (%)
		Weighing day[1]	Weighing day[2]	Weighing day[3]	Overall	
Autumn	Low	821.6	899.9	841.7	854.6 [a]	19.5 [a]
	Medium	1247.1	1234.1	1167.0	1216.1 [b]	30.7 [b]
	High	1864.9	1889	1979.8	1911.2 [c]	36.6 [b]
				SE	*33.1*	
Winter	Low	948.7	917.5	907.2	924.5 [a]	56.4 [c]
	Medium	1285.1	1226.2	1146.2	1219.2 [b]	74.5 [d]
	High	1885.1	1878.4	1847.1	1870.4 [c]	83.0 [d]
				SE	*23.2*	*2.15*
Model effects and comparisons						
Herbage availability level						
High vs Low					***	**
High vs Medium					***	ns
Medium vs Low					***	**
Season (Autumn vs winter)					ns	*
Herbage availability level x Season					ns	*

Herbage availability: Low herbage availability target range: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥1400 kg DM/ha). All tests and comparisons were based on Sidak's multiple comparison methods. Single SEM value for live dry matter comparison across rows and within columns indicates a significant herbage availability x season interaction). *, **,*** indicate significant difference at $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively. ns: indicates not significant ($p > 0.05$).

Appendix IV Figure 1b Herbage quality parameters for grab samples of the Low, Medium, and High herbage availability treatments offered to ewe lambs during autumn and winter (least square means). Analysis conducted using near-infrared reflectance spectroscopy (NIRS) method.

Study season	Herbage availability	Chemical composition				
		DM %	CP %	NDF %	ADF %	ME MJ/kg
Autumn	Low	26.4	16.8	52.9	30.1	9.5
	Medium	26.1	18.3	52	30.2	9.5
	High	26.7	18.4	50.5	28.3	9.8
Winter	Low	19.1	21.6	43.2	23.5	11.5
	Medium	18.7	25.8	42.2	23.1	11.4
	High	19.5	27.3	39.1	21.8	11.4
	SE1	0.58	1.09	1.09	0.76	0.18
	SE2	1.62	1.34	1.33	0.93	0.23
Model effect comparisons						
Herbage availability						
High vs Low		ns	*	ns	ns	ns
High vs Medium		ns	ns	ns	ns	ns
Medium vs Low		ns	*	ns	ns	ns
Season (Autumn vs winter)		*	*	*	*	*
Herbage availability level x Season		ns	ns	ns	ns	ns

DM: dry matter; CP: crude protein; NDF: neutral detergent fiber; ADF: acid detergent fiber (ADF); ME: metabolizable energy. Herbage availability: Low herbage availability target range: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥1400 kg DM/ha. Standard error of mean difference % (*SE1*: comparisons across season; *SE2*: comparisons among herbage levels). * indicates significant difference at $p < 0.05$. ns: indicates not significant ($p > 0.05$).

APPENDIX IV Table 2a Estimated post-feeding herbage mass (least squares means) and proportion of live matter (%) of Low, Medium, and High herbage availability target levels (kg DM/ha) offered to ewe lambs on Tuapaka farm and Riverside farm during validation.

Farm	Herbage availability	Herbage mass (kg DM/ha) Weighing day[1]	Weighing day[2]	Overall	Proportion of live dry matter (%)
Tuapaka	Low	972.3	907.9	940.0	61.7
	Medium	1318.8	1249.7	1284.3	93.4
	High	1921.1	1900.4	1910.8	94.5
Riverside	Low	808.1	956.4	882.3	85.3
	Medium	1277.2	1187.8	1232.5	71.2
	High	1602.6	1458	1530.3	86.2
			SE	*110.0*	*6.2*
Model effects and comparisons					
Herbage availability level					
High vs Low				***	**
High vs Medium				***	ns
Medium vs Low				***	*
Farm (Tuapaka vs Riverside)				*	*
Herbage availability level vs Farm				*	*

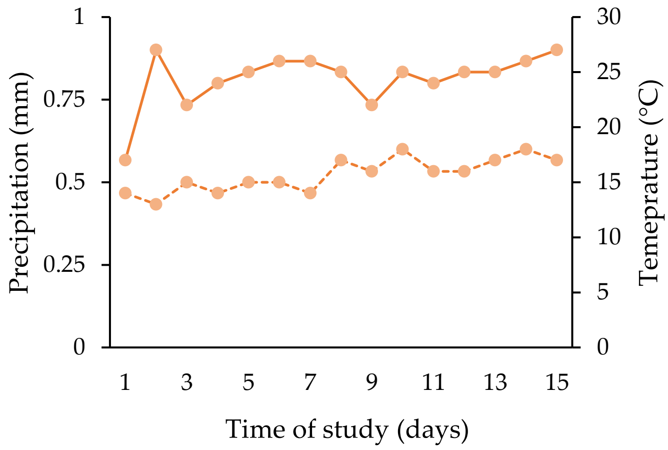
Herbage availability (Low herbage availability target range: 700–900 kg DM/ha, Medium: 1100–1300, High: ≥ 1400). Single SEM value for herbage availability and live matter comparisons across rows and within columns indicates a significant herbage availability level x farm interaction). *, **,*** indicate significant difference at $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively. ns: indicates not significant ($p > 0.05$).

APPENDIX IV Table 2b Herbage quality parameters for grab samples of the Low, Medium, and High herbage availability levels offered to lambs during Winter season on Tuapaka farm and Riverside farm.

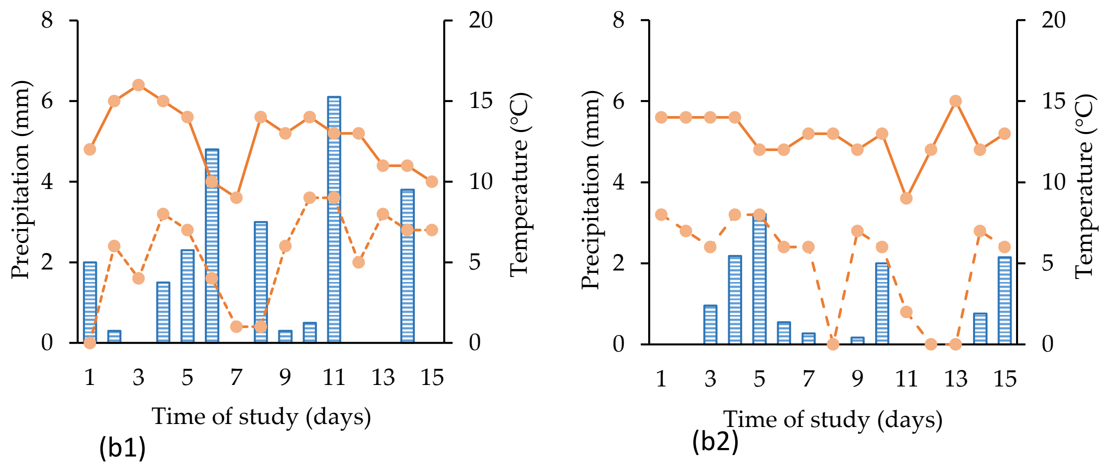
Farm	Herbage availability	Chemical composition DM %	CP %	NDF %	ADF %	ME MJ/kg
Tuapaka	Low	17.9	22.8	39.1	21.8	11.5
	Medium	16.5	26.9	42.2	23.1	11.4
	High	16.7	26.9	43.2	23.5	11.4
Riverside	Low	17.8	24.6	39.1	22.3	10.3
	Medium	18.6	27.2	42.9	19.9	10.8
	High	17.9	27.4	43.3	19.7	10.5
	SE1	0.41	0.46	0.52	0.73	0.06
	SE2	0.4	0.58	0.66	0.81	0.08
Model effect comparisons						
Herbage availability						
High vs Low		*	*	ns	ns	ns
High vs Medium		ns	ns	ns	ns	ns
Medium vs Low		ns	*	*	*	ns
Farm (Tuapaka vs Riverside)		ns	ns	ns	ns	*
Herbage availability level x Farm		ns	ns	ns	ns	ns

DM: dry matter; CP: crude protein, NDF: neutral detergent fiber NDF; ADF: acid detergent fiber; ME: metabolizable energy. Herbage availability (Low herbage availability target range: 700–900 kg DM/ha, Medium: 1100–1300, High: ≥1400). SEM values represent all comparisons across rows. All tests and comparisons were based on Sidak's multiple comparison methods. Standard error of mean difference % (*SE1*: comparison across farm; *SE2*: comparison among herbage levels). * indicates $p < 0.05$, respectively. ns: indicates not significant ($p > 0.05$).

Appendix V: Weather data for Chapter 5

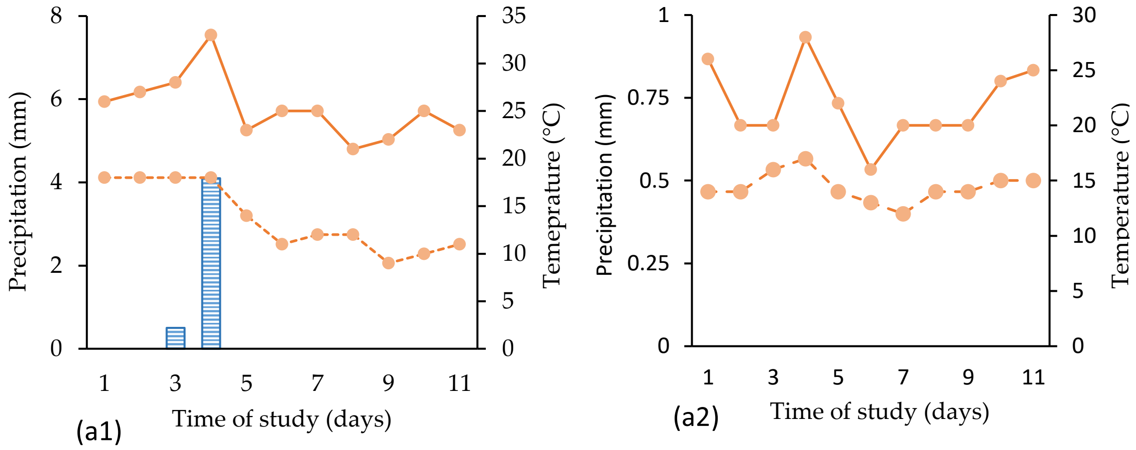


Appendix V Figure 1a Average daily precipitation (stripped bars) and temperature (solid line: maximum, dashed: minimum) during the calibration stage over the study time for non-pregnant ewes. It did not rain during the non-pregnant study.

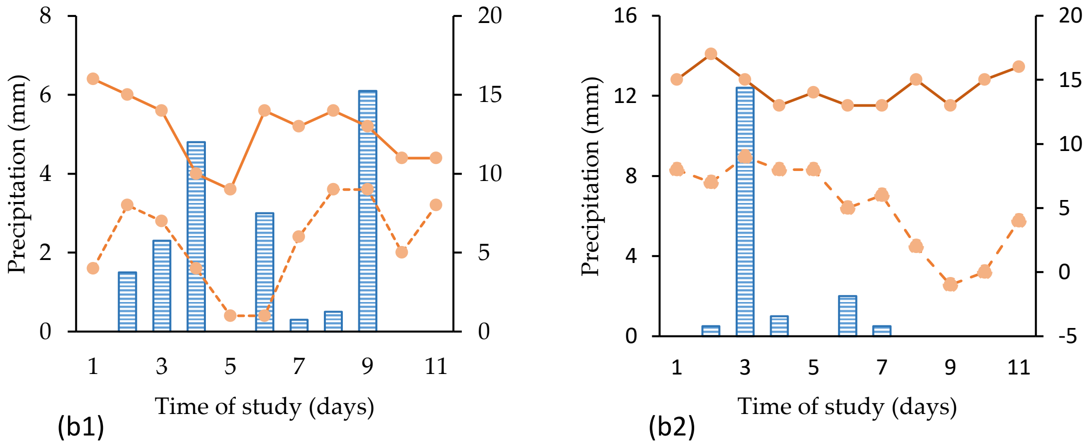


Appendix V Figure 1b Average daily precipitation (stripped bars) and temperature (solid line: maximum, dashed: minimum) during the calibration stage over the study time for the ewes at approximately 100 days from the midpoint of a 17-day breeding period (b1) and 130 days (b2).

Data from: https://cliflo.niwa.co.nz.

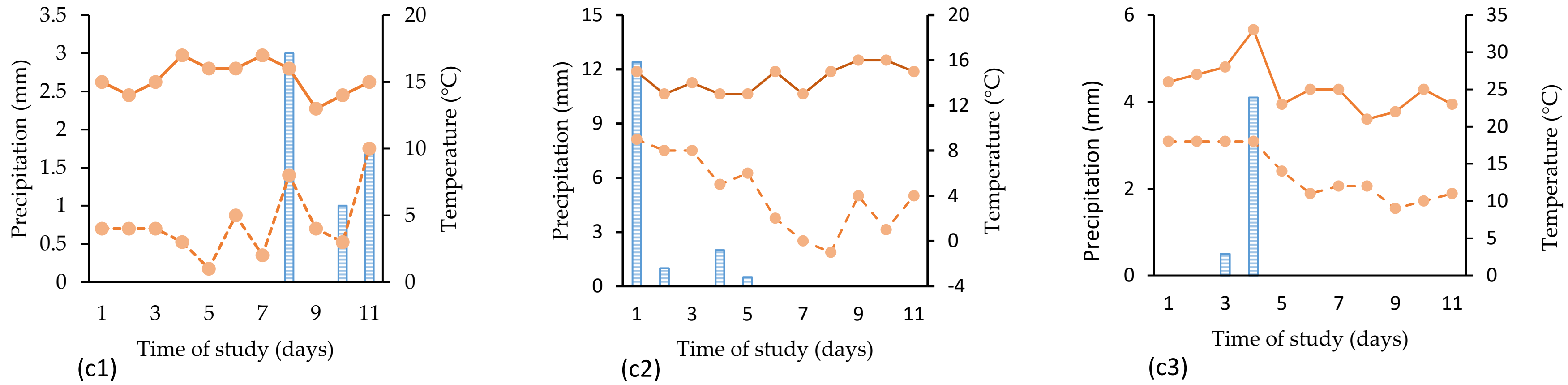


APPENDIX V Figure 2a Average precipitation (stripped bars) and temperature (solid line: maximum, dashed: minimum) during the validation period for non-pregnant ewes at Keeble farm (a1) and Riverside farm (a2).



APPENDIX IV Figure 2b Average precipitation (stripped bars) and temperature (solid line: maximum, dashed: minimum) during the validation period of ewes at approximately 100 days of pregnancy from the midpoint of a 17-day breeding period at Keeble farm (b1) and Tuapaka farm (b2).

Data from: https://cliflo.niwa.co.nz.



APPENDIX IV Figure 3c Average precipitation (stripped bars) and temperature (solid line: maximum, dashed: minimum) during the validation period of ewes at approximately 130 days of pregnancy from the midpoint of a 17-day breeding period at Keeble farm (c1), Tuapaka farm (c2) and Riverside farm (c3).

Data from: https://cliflo.niwa.co.nz.

Appendix VI: Herbage mass and chemical composition for Chapter 5

Appendix VI Table 1a Estimated least squares mean herbage mass (kg DM/ha) and proportion of live/green matter (%) of Low, Medium and High herbage masses offered to ewes by study or physiological state (pregnant, non-pregnant), stage of pregnancy (P100: approximately 100 days of pregnancy from the midpoint of a 17-day breeding period, P130: approximately 130 days) and weighing day (days on which weighing was conducted: 7, 12, 14) during the calibration.

		Herbage mass (kg DM/ha)				Proportion
Study	Herbage availability	Weighing day[1]	Weighing day[2]	Weighing day[3]	Overall	of live dry matter (%)
Non-pregnant ewes						
	Low$_1$	1168.4	1196.4	1185.9	1183. 6	38.6
	Medium	1300.1	1296.1	1210.1	1268.8	67.3
	High	2002	1802	1732.7	1845.6	67.7
				SE	20.81	8.62
Pregnant ewes						
P100	Low$_2$	1025.5	986.8	1037.5	1016.6	80.8
	High	1758.8	1823.5	1602.3	1728.2	75.0
P130	Low$_2$	1057	1083.5	1076.1	1072.2	84.6
	High	1891.4	1839.7	1737.2	1822.8	91.4
				SE	25.9	7.8
Model effects and comparisons						
Non-pregnant ewe study						
Herbage availability level						
High vs Low					***	**
High vs Medium					***	ns
Medium vs Low					***	**
Pregnant ewe study						
Herbage availability (High vs Low)					***	ns
Stage of pregnancy (P100 vs P130)					ns	ns

*,**,*** indicate $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively. ns: indicates not significant ($p > 0.05$).

Appendix VI Table 1b Herbage quality parameters (means with their standard errors in parenthesis) for grab samples of the Low (Low_1, Low_2), Medium and High herbage levels offered to ewes (Least square means) by study (pregnant, non-pregnant ewe study) and stage of pregnancy, during calibration.

Study	Herbage availability	Chemical composition				
		DM %	CP %	NDF %	ADF %	ME MJ/kg
Non-pregnant ewe study	Low_1	47.4	10.3	60.6	33.6	9.3
	Medium	30.4	12.2	54.3	32.4	8.9
	High	34.4	11.6	56.1	32.0	9.1
	SE	2.21	2.27	3.22	2.52	0.54
Pregnant ewe study						
P100	Low_2	19.6	24.0	48.3	23.8	10.3
	High	15.7	24.4	39.1	19.7	10.5
P130	Low_2	18.8	21.2	41.7	20.0	10.9
	High	15.9	24.1	38.7	18.2	11.7
	SE	0.57	0.81	1.70	1.10	0.24
Model effect comparisons						
Non-pregnant ewe study						
Herbage availability						
High vs Low		*	ns	ns	ns	ns
High vs Medium		ns	ns	ns	ns	ns
Medium vs Low		*	ns	ns	ns	ns
Pregnant ewe study						
Herbage availability (High vs Low)		ns	ns	*	*	*
Stage of pregnancy (P100 vs P130)		ns	ns	ns	ns	ns

DM: dry matter, neutral detergent fibre (NDF), ADF: acid detergent fibre (ADF), CP: crude protein, ME: metabolizable energy. Herbage availability (For non-pregnant ewe study, Low_1 target range herbage mass of: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥1400 kg DM/ha; for pregnant ewe study, Low_2 herbage target range : 900–1100 kg DM/ha, High: ≥1400 kg DM/ha). *,**,*** indicate $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively. ns: indicates not significant ($p > 0.05$). Means comparisons were based on Sidak's adjustment method.

Appendix VI Table 2a Estimated herbage mass (least squares means) and proportion of live and dead matter of Low, Medium and High herbage levels (kg DM/ha) offered to ewe by physiological state (pregnant, non-pregnant) and day of ewe weighing at different farms (Keeble, Tuapaka, Riverside) during the validation.

			Herbage mass (kg DM/ha)			
Study	Farm	Herbage availability	*Weighing day one	**Weighing day two	Overall	Live/green matter (%)
Non-pregnant ewes						
	Keeble	Low_1	1225.4	1298.1	1261.75	56.7
		Medium	1309.2	1238.9	1274.05	71.8
		High	2002.0	1832.6	1917.3	74.7
		SEM[1]			*63.2*	
	Riverside	Low_1	-	-		-
		Medium	1147.9	1116.3	1132.1	0
		High	-	-		-
Pregnant ewes						
P100	Keeble	Low_2	1017.5	930.0	973.7	72.3
		High	2141.0	2026.9	2083.9	96.2
	Tuapaka	Low_2	986.8	967.9	977.4	82.7
		High	1923.5	1711.9	1817.7	88.3
		SEM[2]			*94.4*	
P130	Keeble	Low_2	1105.9	990.0	1047.9	84.6
		High	1892.9	1815.1	1854.0	91.4
	Tuapaka	Low_2	1040.0	1030.4	1035.2	75.4
		High	2169.5	1727.3	1948.4	84.9
	Riverside	Low_2	1081.6	990.0	1035.8	78.8
		High	1681.5	1564.1	1622.8	81.4
		SEM[3]			*(59.4,72.3)*	
Model effect comparisons						
Non-pregnant ewe study						
Herbage availability						
High vs Low					***	**
High vs Medium					***	Ns
Medium vs Low					ns	**
Pregnant ewe study						
Herbage availability (High vs Low)					**	*
†Pregnancy stage (P100 vs P130)					ns	Ns
Farm					*	Ns
Farm x Herbage availability						
P100					**	Ns
P130					.	Ns

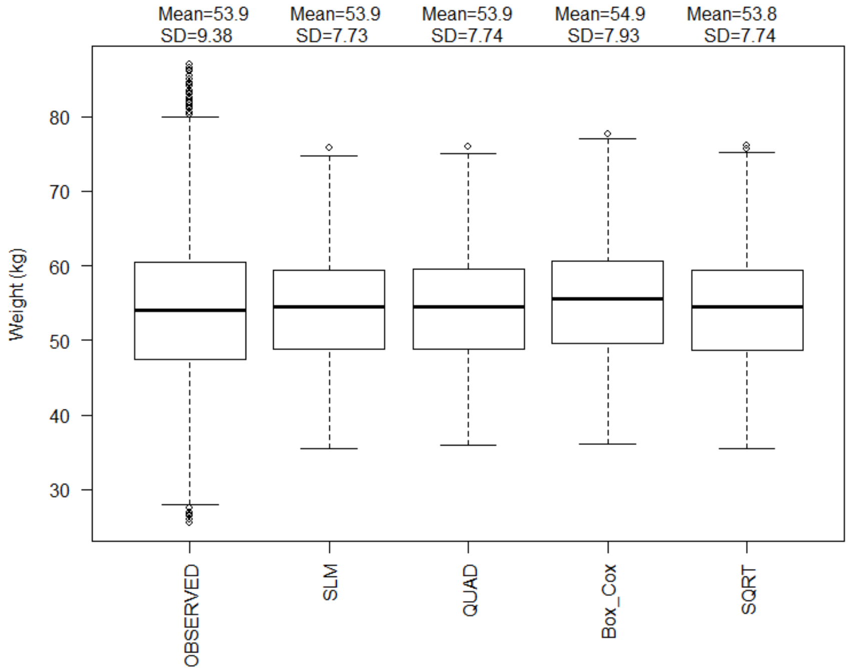
Herbage availability (For non-pregnant ewes, Low[1] target range: 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥1400 kg DM/ha; for pregnant ewes, Low_2 herbage target range: 900–1100 kg DM/ha, High: ≥ 1400;). Stage of pregnancy (P100: 100 days of pregnancy from the midpoint of a 17-day breeding period, P130: 130 days). - Indicates study not conducted. ·, *,**,*** indicate marginally significant ($p \geq 0.05$), significant at $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively. ns indicates not significant. † Two stages within the pregnant ewe study. SEM (1: one-way herbage availability; 2: two-way herbage x farm; 3: no interaction two factors (herbage availability = 59.4, farm = 72.3).

APPENDIX VI Table 2b Herbage quality parameters for hand-plucked samples of the Low, Medium and High herbage levels by study (pregnant, non-pregnant ewe study), farm (Keeble, Tuapaka and Riverside), and stage of pregnancy (P100: 100 days of pregnancy from the midpoint of a 17-day breeding period, P130: 130 days) offered to ewes pre-fasting during validation.

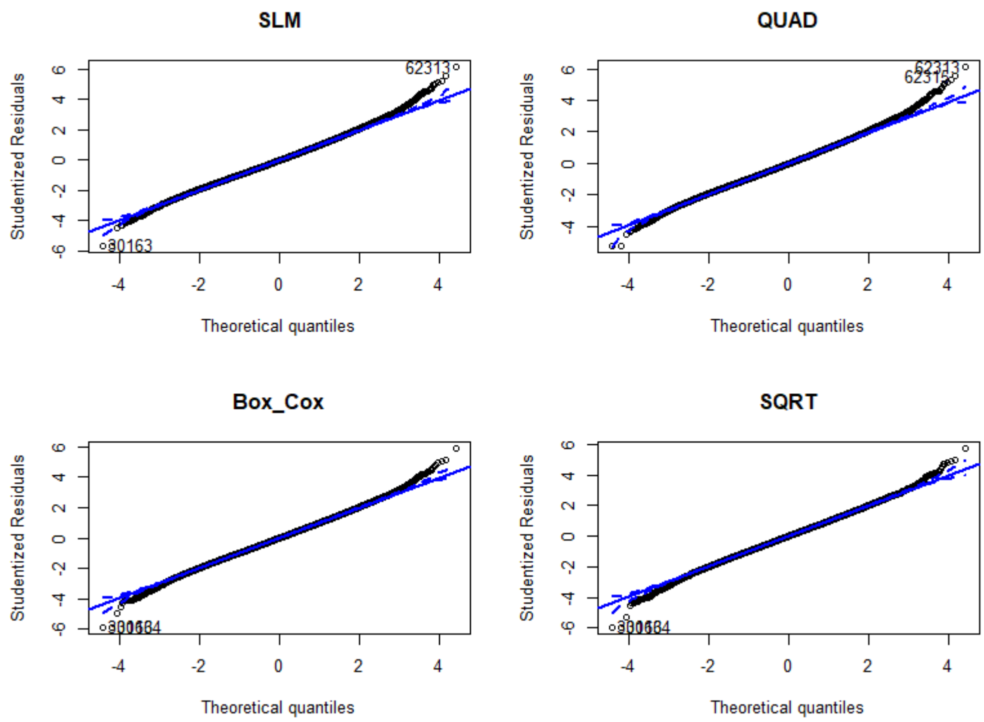
Study	Farm	Herbage availability	DM %	CP %	NDF %	ADF %	ME %
Non-pregnant ewes							
	Keeble	Low_1	33.7	12.2	60.1	33.7	8.9
		Medium	27.7	14.1	53.2	32.2	9.1
		High	25.5	13.1	51.7	30.4	9.3
		SEM	2.29	0.75	1.83	1.03	0.26
	Riverside	Low_1	-	-	-	-	-
		Medium	87.2	7.1	66.2	36.4	8.4
		High	-	-	-	-	-
Pregnancy ewes							
P100	Keeble	Low_2	19.8	20.0	51.8	30.6	9.4
		High	16.0	24.4	37.6	20.6	11.0
	Tuapaka	Low_2	16.3	24.0	41.7	24.6	10.9
		High	13.4	27.5	38.9	20.0	12.2
		SEM	0.68	1.82	2.20	1.60	0.41
P130	Keeble	Low_2	20.9	24.0	39.5	19.9	11.5
		High	14.4	29.9	41.8	20.2	11.7
	Tuapaka	Low_2	15.7	22.2	45.3	26.0	10.3
		High	12.7	29.6	37.8	18.2	12.0
	Riverside	Low_2	18.6	16.	50.4	28.0	10.2
		High	18.8	18.9	48.4	27.4	9.7
		SEM	0.80	1.70	2.16	1.56	0.33
Model effects and comparisons							
Herbage availability (non-pregnant vs pregnant ewe study)			**	**	*	*	*
Non-pregnant ewe study							
Herbage availability							
High vs Low			ns	ns	*	ns	ns
High vs Medium			ns	ns	ns	ns	ns
Medium vs Low			ns	ns	*	ns	ns
Pregnant ewe study							
Herbage availability (High vs Low)			**	ns	*	*	ns
†Pregnancy stage (P100 vs P130)			**	ns	ns	ns	ns
Farm			*	ns	*	*	ns
Farm x Herbage availability			*	ns	*	*	ns

DM: Dry matter, CP: Crude protein, Neutral detergent fibre (NDF), ADF: acid detergent fibre (ADF), ME: metabolizable energy. Herbage availability (Non-pregnant ewes, Low_1 target range herbage : 700–900 kg DM/ha, Medium: 1100–1300 kg DM/ha, High: ≥1400 kg DM/ha; for pregnant ewes, Low_2 target range herbage: 900–1100 kg DM/ha, High (H): ≥1400). -indicates data not collected. *,**,*** indicate $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively. † Two stages within the pregnant ewe study.

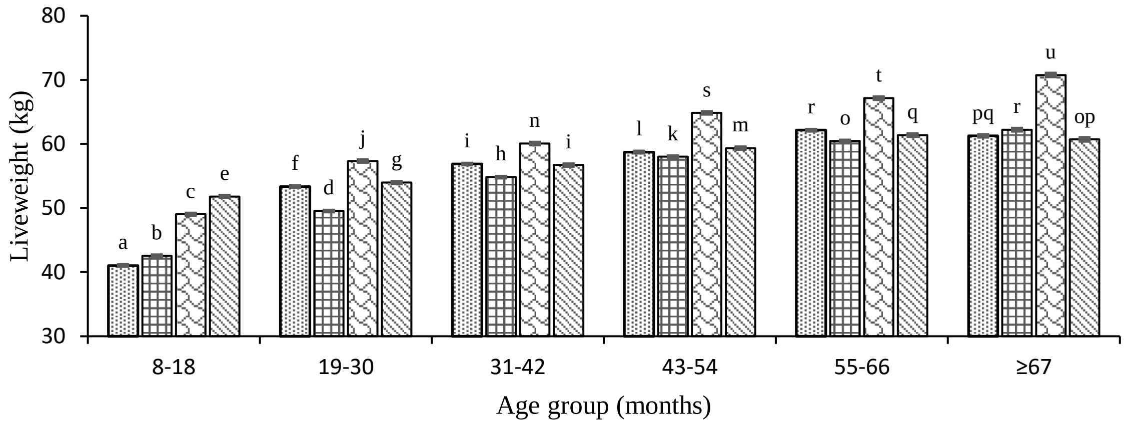
Appendix VII: Box and residual plots, liveweight and body condition score trends (Chapter 6)



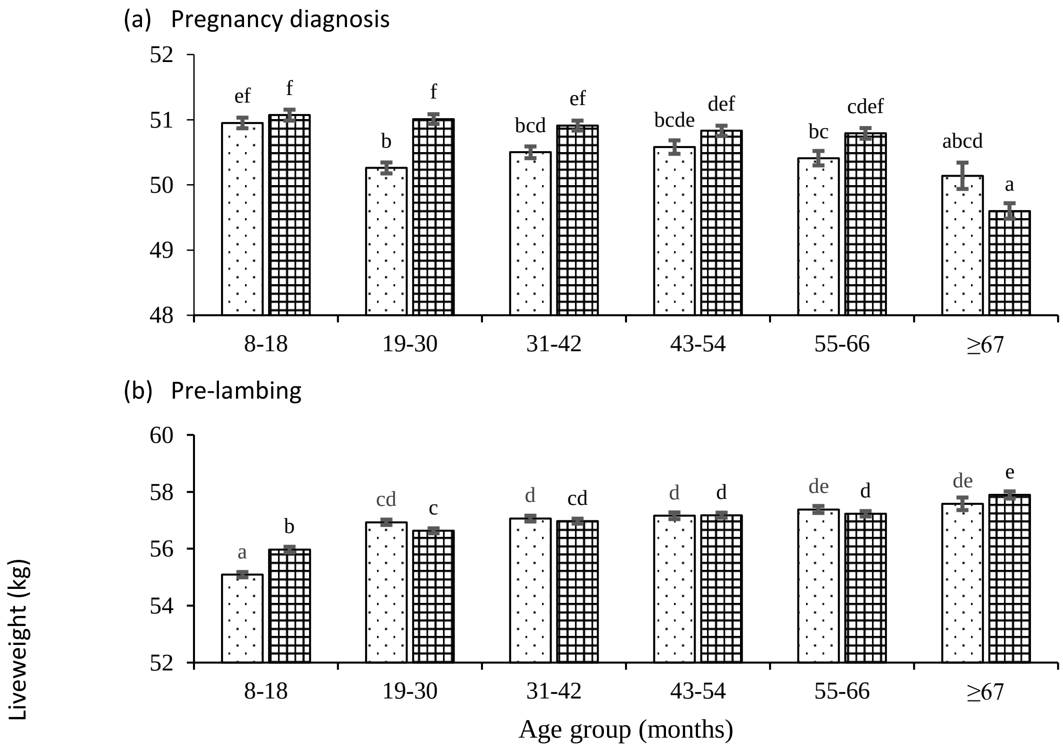
Appendix VII Figure 1a Boxplot summarizing liveweight predictions on the testing dataset; the bottom and top of the box show the 25th and 75th percentiles respectively, whiskers present the 1.5 times the interquartile range of the data, and the thick black solid horizontal line is the median. SD is the standard deviation in kg.



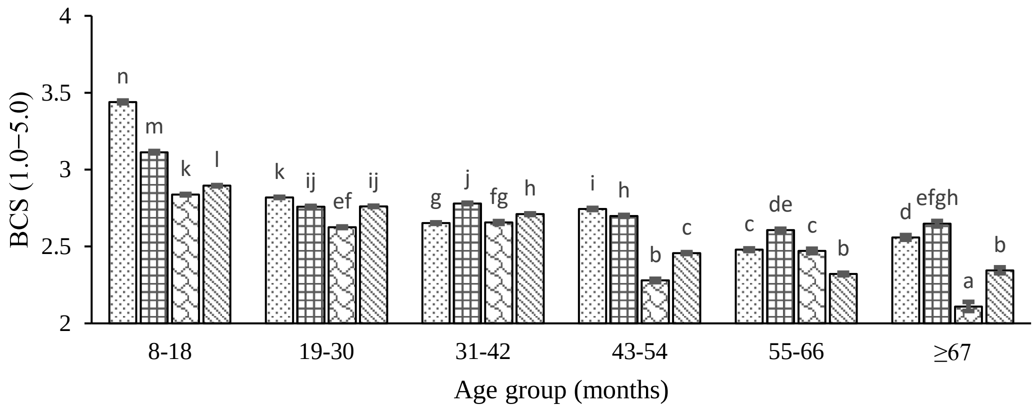
Appendix VII Figure 1b Normal quantile plots of the studentized residuals for all four models (SLM, QUAD, Box_Cox, SQRT) using BCS (1.0–5.0) to predict liveweight (kg). Solid blue line indicates simulated robust regression. Dotted lines indicate the confidence envelope estimated by parametric bootstrap (repeats =1000).



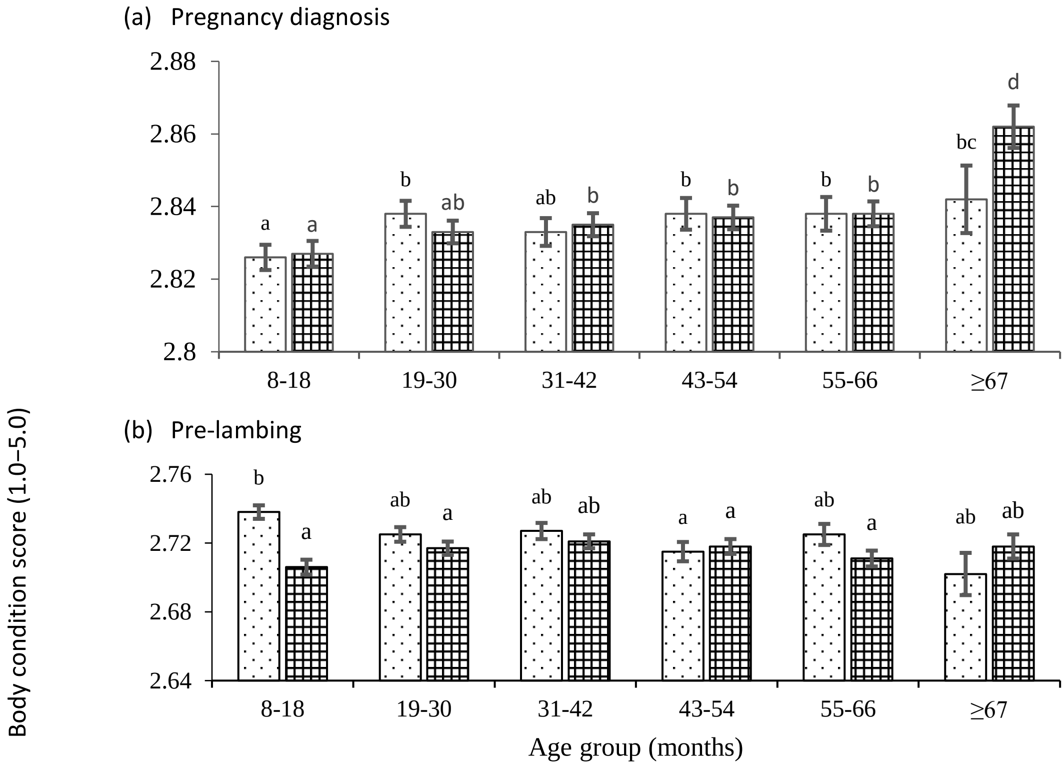
Appendix VII Figure 2a The liveweight of ewes across each age group (8−18 months, 19−30, 31−42, 43−54, 55−66 and ≥67) and stage of the annual cycle (dotted bar: pre-breeding, grid: at pregnancy diagnosis, shingled: pre-lambing and stripped-diagonal: weaning). Superscripts a - u indicate significant differences ($p < 0.05$) across age group and stage of the annual cycle.



Appendix VII Figure 2b The liveweight of ewes across each age group (8−18 months, 19−30, 31−42, 43−54, 55−66 and ≥67) and pregnancy diagnosis (dotted bar: non-pregnant, grid: single, shingled: twin). Superscripts a - f indicate significant differences ($p < 0.05$) across age group and stage of the annual cycle.



Appendix VII Figure 3a The BCS of ewes across each age group (8−18 months, 19−30, 31−42, 43−54, 55−66 and ≥67) and stage of the annual cycle (dotted bar: pre-breeding, grid: at pregnancy diagnosis, shingled: pre-lambing and stripped-diagonal: weaning). Superscripts a - n indicate significant differences ($p < 0.05$) across age group and stage of the annual cycle.



Appendix VII Figure 3b The liveweight of ewes across each age group (8−18 months, 19−30, 31−42, 43−54, 55−66 and ≥67) and pregnancy diagnosis (dotted bar: non-pregnant, grid: single, shingled: twin). Superscript a-d indicates significant differences ($p < 0.05$) across age group and pregnancy-rank.

Appendix VIII: Regression coefficients for BCS prediction equations in Chapter 7

Appendix VIII Table 1a Linear regression intercepts and coefficients and adjusted R^2 for the prediction of BCS from liveweight (liveweight alone models) between 8–18 and 32–43 months of ewes age across stages of reproductive cycle.

Predictor	BM1	BP1	BL1	BW1	BM2	BP2	BL2	BW2	BM3	BP3	BL3	BW3
WM1	0.04	−0.02	−0.01	−0.02	−0.01	−0.01	−0.02		−0.01	−0.01	−0.01	−0.01
WP1		0.04	0.02		0.01	−0.03	−0.03	−0.03	−0.01	−0.01		0.01
WL1				−0.01	−0.02	0.02	0.02	0.02				−0.01
WW1				0.05	0.01		0.01	−0.01		−0.01	−0.01	
WM2					0.03	0.01	0.01			−0.01		−0.01
WP2						0.03		−0.01	−0.01		0.01	0.01
WL2							0.01				−0.01	−0.01
WW2								0.05	0.01			
WM3									0.04	0.01	0.01	−0.01
WP3										0.05	0.01	−0.01
WL3											0.02	
WW3												0.05
WM4												
WP4												
WL4												
WW4												
WM5												
WP5												
WL5												
WW5												
WM6												
WP6												
WL6												
WW6												
Intercept	1.40	2.14	2.6	1.27	1.33	1.62	2.26	1.26	1.69	1.26	1.94	1.84
Adjusted R^2	14.1	8.19	6.2	45.4	38.4	25.5	24.8	36.4	38.7	38	14.9	48.9

BM, BP, BL, BW indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively. WM, WP, WL, WW indicate liveweight prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively. Blank space indicates coefficient non-significant at $p < 0.05$. Model example for BCS estimation (e.g., BM1 = 1.41 + 0.04 WM1, adj. R^2 = 14%).

Appendix VIII Table 1b Linear regression intercepts and coefficients and adjusted R^2 for the prediction of ewe BCS from liveweight (liveweight alone models) above 43 months of age across stages of reproductive cycle.

Predictor	BM4	BP4	BL4	BW4	BM5	BP5	BL5	BW5	BM6	BP6	BL6	BW6
WM1	−0.01	−0.01				−0.01	−0.01	−0.01	−0.01	−0.02	−0.01	−0.02
WP1	−0.01	−0.01	0.01		0.01				−0.01	−0.01		
WL1			−0.02	−0.01	−0.01					0.01		0.01
WW1			−0.02	−0.01	−0.01	−0.01				−0.01	−0.01	−0.01
WM2			−0.01	−0.01	−0.01	−0.01		−0.01		0.01		0.01
WP2	−0.01		0.02	0.01	0.01				−0.01		0.01	−0.02
WL2					0.01	0.01						0.01
WW2	−0.01		0.01		0.01					−0.01		
WM3		−0.01	−0.01	−0.01	−0.01	−0.01			−0.01	−0.01	0.01	
WP3						0.01					−0.01	
WL3						−0.01				0.01		
WW3	0.02		0.01			−0.01			−0.01		−0.01	
WM4	0.03									−0.01		
WP4		0.04		−0.01		0.01			0.01	0.02	0.01	
WL4			0.02							−0.01	−0.01	−0.01
WW4				0.04	0.01							
WM5					0.03	0.01	0.01	−0.01			0.01	0.01
WP5						0.03	0.01			0.01	0.01	−0.01
WL5							0.01	−0.01	−0.01	−0.01	−0.02	−0.01
WW5								0.05	0.01	0.01		
WM6									0.04	0.01	0.01	
WP6										0.03		−0.01
WL6											0.02	
WW6												0.06
Intercept	1.59	2.30	2.40	1.72	1.46	1.59	1.92	1.65	1.71	1.60	1.96	1.05
Adjusted R^2	44.7	32.35	48.9	41.66	36.6	28.01	14.8	52.86	52.59	39.27	11.6	46.94

BM, BP, BL, BW indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively. WM, WP, WL, WW indicate liveweight prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively. Blank space indicates coefficient non-significant at $p < 0.05$. Model example for BCS estimation (e.g., BM4 = 1.59–0.01 WM1+ … + 0.03 WM4, adj. R^2 = 45%).

Appendix VIII Table 2a Linear regression intercepts and coefficients and adjusted R^2 for the prediction of ewe BCS from combined models (that included lifetime liveweight, liveweight change, and previous BCS) between 8–18 and 32–43 months of ewe age across stages of reproductive cycle.

Predictor	BM1	BP1	BL1	BW1	BM2	BP2	BL2	BW2	BM3	BP3	BL3	BW3
WM1	0.04			−0.01		−0.01	−0.03		−0.01		−0.01	−0.01
BM1		0.16	0.015	0.016	0.011	0.018	0.09	0.011	0.06	0.08	0.07	0.01
WP1				−0.02		−0.02	0.02			−0.01		
DWT11		0.11	0.01	0.01			−0.02		−0.01			−0.01
BP1			0.04	0.07	0.011	−0.014	−0.04	−0.019	0.08	0.01	0.07	0.013
DWT12			0.01				0.01					−0.01
WL1				0.04		−0.01	−0.01	−0.01			0.01	
BL1				0.01	0.09	0.012	0.04	0.05	0.05	0.01	0.04	0.08
DWT13				0.05	0.02	−0.02	−0.01	−0.01			0.01	
WW1				0.01		0.01	0.02			−0.01	−0.02	
BW1					0.028	0.08	0.05	0.07	0.03	−0.01	0.03	0.02
DT2-T1					0.02							
WM2					0.02				−0.01	−0.01	−0.04	−0.01
BM2						0.013	−0.03	0.08	0.09	0.09	−0.01	0.09
DWT21							−0.01				−0.04	
WP2						0.03			−0.01		0.02	−0.01
BP2							0.051	0.024	0.01	0.013	0.01	0.03
DWT22								0.02			−0.02	−0.02
WL2								0.07		0.01	0.02	0.02
BL2								0.011	0.09	0.07	0.015	−0.07
DWT23								0.09		0.01		
WW2								−0.04		−0.02	−0.02	
BW2									0.023	0.018	0.03	0.04
DT3-T2											−0.01	
WM4									0.03		−0.02	−0.01
BM3										0.022	0.011	0.01
DWT31											−0.03	
WP3										0.04	0.08	−0.02
BP3											0.036	0.06
DWT32											0.04	
WL3											−0.03	
BL3												0.025
DWT33												0.01
WW3												0.05
Intercept	1.40	2.30	1.20	0.83	0.42	1.12	1.90	0.65	0.52	0.10	0.22	0.30
Adjusted R^2	14.10	10.5	34.0	51.0	50.33	32.0	43.55	58.0	54.02	55.43	33.48	56.52

BM, BP, BL, BW indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively. WM, WP, WL, WW indicate liveweight prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively. DWT, DW-T indicate liveweight change within age group and between age groups, respectively. Blank space indicates coefficient non-significant at $p < 0.05$. Model example for BCS estimation (e.g., BM1 = 1.41 + 0.04 WM1, adj. R^2 = 14%).

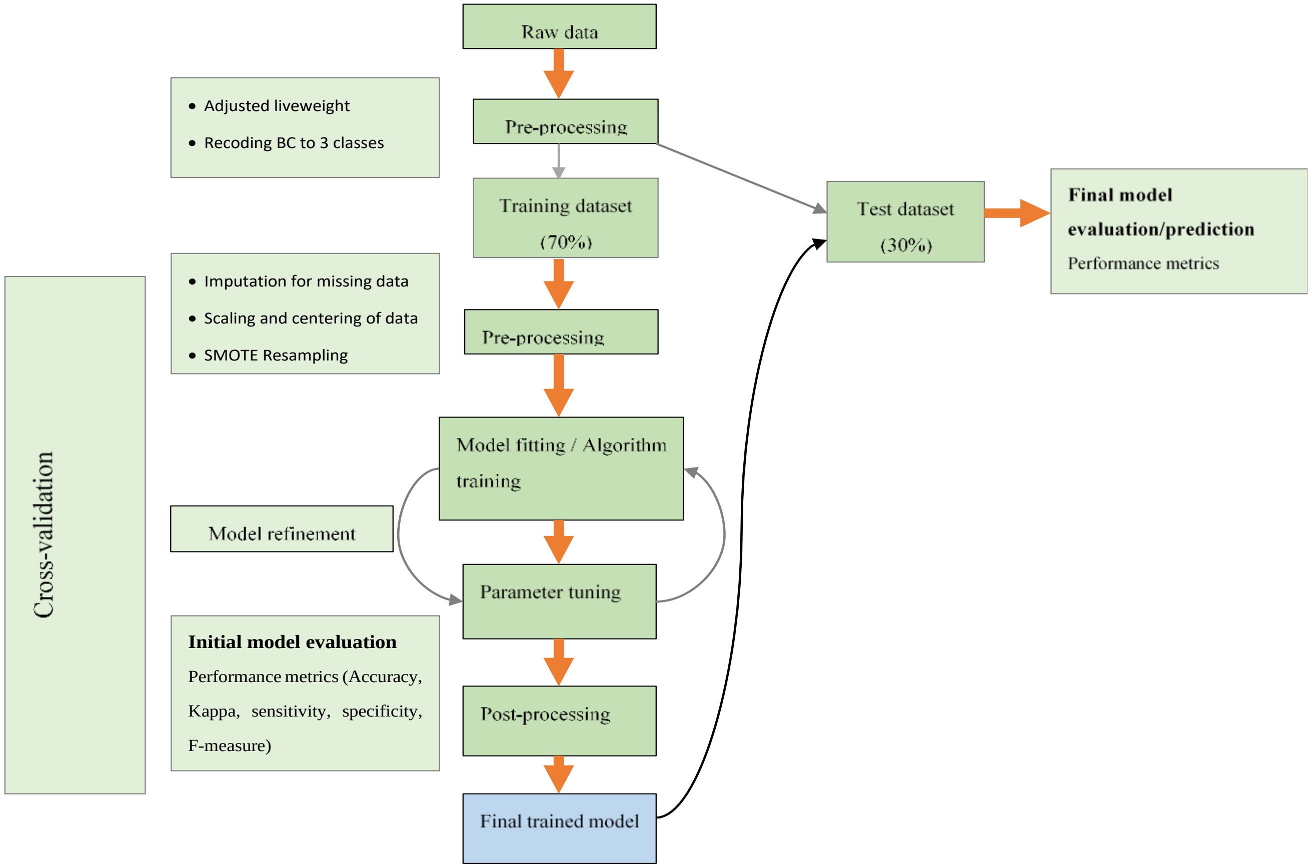
Appendix VIII Table 2b Linear regression intercepts and coefficients and adjusted R^2 for the prediction of ewe BCS from combined models (that included lifetime liveweight, liveweight change, and previous BCS) above 43 months of ewe age across stages of reproductive cycle.

Predictor	BM4	BP4	BL4	BW4	BM5	BP5	BL5	BW5	BM6	BP6	BL6	BW6
WM1			0.01		0.01				0.02			−0.02
BM1	0.06	0.01		0.03	−0.01	0.03	0.02	−0.03	0.02	0.04	0.03	0.07
WP1	−0.01	−0.01							−0.02		−0.01	
DWT11	0.01								0.01	0.01		
BP1	0.05	0.01	−0.08	0.03	0.03	0.01	0.02	0.01	0.04	−0.01	−0.02	−0.01
DWT12		−0.01	−0.02		−0.01	−0.01				0.01	−0.01	
WL1				−0.01	−0.01		−0.01	0.01	0.01	0.01	0.01	0.01
BL1		0.06	0.01	0.08	0.03		0.09	0.05	0.02	−0.02	0.09	0.01
DWT13	−0.01			−0.01			−0.01			0.01	0.01	0.01
WW1		−0.01	−0.01		−0.01	−0.01	0.01	−0.01				0.02
BW1	0.05		−0.07	−0.01	0.04	0.02		0.03	−0.02	0.01	0.04	−0.02
DT2-T1	−0.01	−0.02							−0.01	0.02	0.01	0.03
WM2		0.01	−0.01			−0.01	−0.02	−0.01		−0.01	0.02	−0.06
BM2	0.02	0.04	0.01	0.04	0.06	0.03	0.03	0.02	0.01	0.01	0.04	0.07
DWT21		−0.01				−0.01	−0.01	−0.01			0.03	−0.02
WP2	−0.01	−0.01	0.03	0.01		0.03	0.02	0.01	−0.02	−0.01		
BP2	0.01	0.07	0.04	0.04	0.02	0.05	0.04	0.06	0.08	0.05	−0.01	−0.05
DWT22		−0.02	0.02			0.02	0.01	−0.01	−0.01	−0.01	0.03	−0.02
WL2	−0.01		−0.02	0.01	0.01		−0.02	0.01	0.01	0.01	−0.02	0.05
BL2	0.04	0.05	0.03	−0.02		0.08	0.07	−0.01	0.05	0.05	0.01	0.02
DWT23	−0.01	−0.01		0.01	0.01	0.01	−0.01		−0.01		0.01	0.02
WW2			0.01	−0.01	−0.01	−0.02		−0.01		−0.02	−0.05	−0.01
BW2	0.03		0.01	0.09	0.04	0.04	0.03	0.06	0.06	0.05	−0.03	0.08
DT3-T2			0.01					−0.01			−0.04	0.02
WM3	−0.01		−0.02	0.01	−0.01	−0.01	−0.01		−0.01			0.01
BM3	0.02	0.01	0.06	0.07	0.08	0.02	0.08	0.07	0.08	0.04	0.03	0.08
DWT31		0.01		0.02		−0.01	−0.01	−0.01			−0.05	0.03
WP3	−0.01	−0.02	−0.01	−0.02		0.01		0.01		0.01		−0.04
BP3	0.01	0.08	0.08	0.03	0.04	0.09	0.08	0.02	0.07	0.07	−0.01	0.01
DWT32	0.01	−0.01	−0.01		0.01		−0.01	0.01			−0.05	−0.02
WL3			−0.01					−0.01	−0.02		0.04	0.05
BL3	0.01	0.01	0.06		0.02	0.06	0.04	0.03	0.01	0.05	0.01	−0.05
DWT33			−0.02				−0.01		−0.02	0.01		0.04
WW3	0.01	−0.01	0.02	−0.01	−0.01			−0.01	0.01	−0.01	−0.01	−0.03
BW3	0.02	0.01	0.09	0.01	0.01	−0.01	0.05	0.04	−0.02	−0.04	0.01	0.04
DT4-T3						−0.01	−0.01			−0.01		
WM4	0.03	0.01	−0.05	−0.01	−0.01		0.01		0.01		0.03	
BM4		0.02	0.03	0.04	0.06	0.04		0.01	0.08	0.06	0.04	0.01
DWT41		0.01	−0.05	−0.01	−0.01	0.01	0.01	0.01	0.01	0.01	0.02	−0.01
WP4		0.03	0.04	−0.02		−0.01	−0.02	−0.01				−0.03
BP4			0.01	0.08	0.01	0.01	0.08	0.08	0.01	0.02	−0.04	0.03
DWT42			−0.01	−0.01	−0.01				0.01	0.01	0.03	−0.01
WL4			0.03		0.01			0.01			−0.03	0.02
BL4				0.02	0.01	0.01	0.08	0.06	−0.01	0.01	0.01	0.01
DWT43				−0.01				0.01	0.01	0.02		0.02
WW4				0.05	0.01			−0.01	−0.01	−0.02	−0.01	−0.01
BW4					0.02	0.08	0.07	0.09	0.04	−0.04	−0.07	−0.03
DT5-T4									0.01	−0.01	−0.01	
WM5					0.04	0.09			0.01	0.01		0.02
BM5						0.01	0.02	0.04	0.08	0.06	0.03	0.03
DWT51						0.09	−0.01		0.02		−0.01	0.02
WP5						−0.06	0.03	−0.01	−0.03	−0.02	−0.01	−0.05

BM5							0.07	0.08	0.05	0.02	0.03	0.05
DWT52							0.01				−0.02	−0.03
WL5							0.01	−0.01	0.02		0.01	
BL5								0.03	0.08	0.01	0.05	0.01
DWT53								−0.01	0.03			−0.04
WW5								0.07	−0.02			0.02
BW5									0.02	0.08	0.01	0.07
DT6-T5											−0.01	−0.01
WM6									0.05		0.02	0.01
BM6										0.02	0.01	0.01
DWT61										−0.01	0.01	
WP6										0.04	0.01	0.01
BP6											0.01	0.02
DWT62											0.01	0.03
WL6												0.01
BL6												0.02
DWT63												0.04
WW6												0.02
Intercept	0.02	0.03	0.07	0.05	0.03	0.02	0.04	0.01	0.05	0.04	0.05	0.14
Adjusted R^2	53.98	47.73	51.48	50.88	48.92	46.22	31.43	59.05	61.4	57.96	24.19	49.67

BM, BP, BL, BW indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively. WM, WP, WL, WW indicate liveweight prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively. DWT, DW-T indicate liveweight change within age group and between age groups, respectively. Blank space indicates coefficient non-significant at $p < 0.05$. Model example for BCS estimation (e.g., BM4 = 0.02 + 0.06 BM1+ … +0.03 WM4, adj. R^2 = 54%).

Appendix IX: Machine learning flow chart, model comparison and summary of indicators of accuracy (Chapter 8)



Appendix IX Figure 1 Machine learning flow chart for ewe BCS prediction using their current and previous liveweights.

Appendix IX Table 1 Accuracy measures (Precision, F-measure) of nine predictive models for ewe BCS at 43−54 months of age pre-breeding at different stages of the annual sheep weighing cycle (PB: pre-breeding, PD: pregnancy diagnosis, PL: pre-lambing and W: weaning). Values in parenthesis indicate the minimum and maximum.

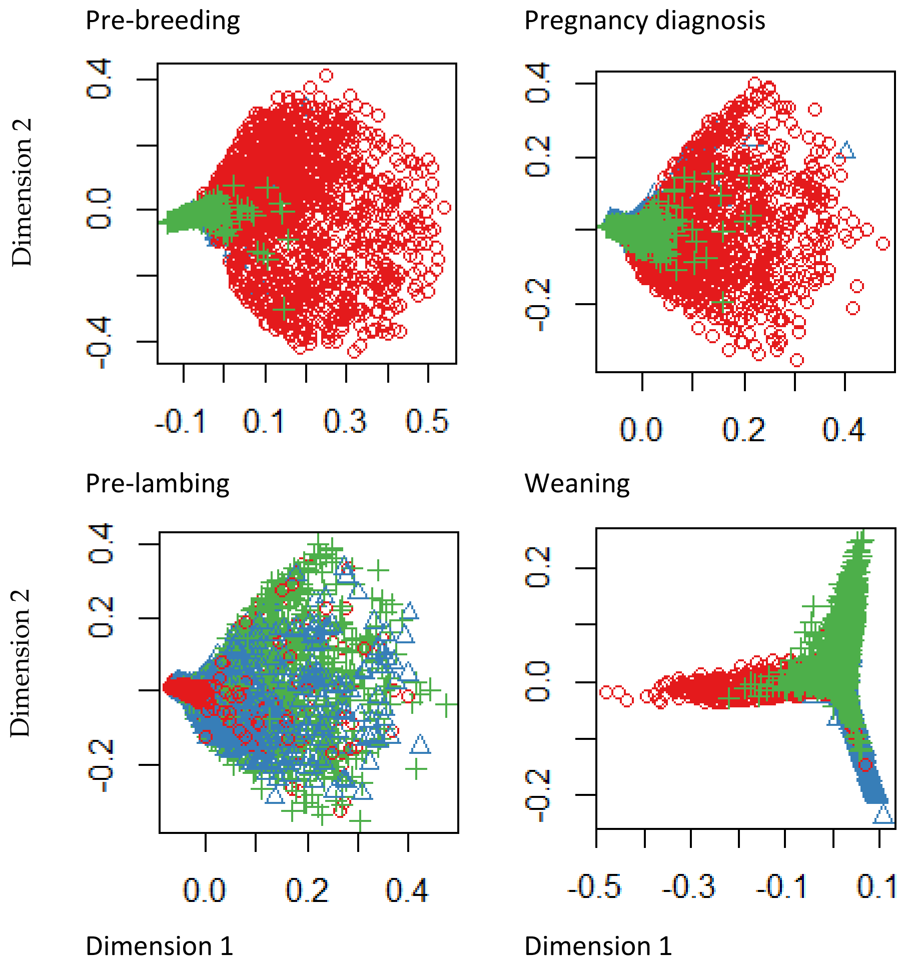
	PB		PD		PL		W	
Model	Precision %	F-measure %	Precision %	F-measure %	Precision %	F-measure %	Precision %	F-measure %
XGB	86.1 (78.2-97.7)	86.0 (80.1-96.9)	87.9 (80.8-94.5)	87.6 (84.1-90.0)	87.9 (80.8-94.5)	87.6 (84.1-90.0)	89.1 (84.2-92.8)	89.0 (87.5-91.3)
RF	85.3 (78.1-95.9)	85.3 (79.0-95.6)	86.9 (83.2-91.1)	86.7 (83.6-90.7)	86.1 (77.0-91.7)	85.7 (81.0-89.0)	84.9 (79.3-88.8)	84.7 (83.2-86.4)
SVM	82.7 (74.1-95.1)	82.7 (74.5-94.4)	83.4 (74.6-90.3)	82.6 (80.0-87.2)	83.5 (68.7-95.0)	81.8 (76.0-86.4)	82.8 (71.6-89.4)	82.0 (78.0-85.7)
KNN	82.3 (75.0-94.4)	82.0 (71.8-95.3)	84.7 (77.5-89.5)	84.5 (80.9-90.6)	64.5 (58.1-68.6)	64.1 (61.8-65.5)	84.9 (79.3-88.8)	85.1 (80.5-88.1)
ANN	80.3 (71.9-93.4)	80.3 (71.6-92.6)	76.3 (72.1-83.7)	76.1 (73.2-80.7)	73.5 (64.5-83.3)	71.5 (67.4-76.2)	79.5 (70.0-85.0)	78.7 (76.4-82.6)
Multinom	76.8 (67.7-89.3)	76.8 (68.1-89.1)	70.2 (65.6-76.4)	70.0 (64.1-73.8)	64.8 (62.8-65.9)	64.6 (62.1-67.1)	68.1 (65.0-70.7)	67.7 (65.7-70.2)
LDA	75.0 (64.3-89.0)	74.9 (64.5-88.3)	70.5 (65.1-79.0)	69.3 (61.8-73.3)	65.3 (61.9-67.9)	64.9 (61.5-67.7)	68.3 (63.4-70.8)	67.6 (65.8-70.7)
Ordinal	73.2 (59.2-88.5)	72.9 (60.4-85.3)	64.9 (55.0-77.4)	63.8 (58.4-68.1)	57.3 (45.8-64.2)	57.5 (43.5-66.7)	64.2 (52.9-70.9)	63.4 (57.4-68.7)
CART	62.1 (47.3-77.7)	62.3 (41.5-80.0)	61.1 (55.5-68.9)	59.2 (48.5-64.6)	57.3 (55.1-60.5)	55.7 (46.6-62.5)	55.4 (53.4-59.0)	54.8 (45.3-60.9)

Model: (XGB: Gradient boosting decision tree model, RF: Random Forest, KNN: K-Nearest Neighbours, SVM: Support Vector Machines, ANN: Neural Networks, Multinorm: Multinomial regression, LDA: Linear Discriminant Analysis, Ordinal: ordinal logistic regression, CART: Classification and regression tree).

Appendix IX Table 2 A pairwise comparison (Bonferroni *p*-value adjustment) of overall performance accuracy of nine predictive models for BCS, at different stages of the annual cycle (PB: pre-breeding, PD: pregnancy diagnosis, PL: pre-lambing, W: weaning) in 43−54-month-old ewes. *p*-value > 0.05 indicates not significant difference between models. All ewe BCS predictions were based on liveweight records.

Model A	Model B	PB	PD	PL	W
XGB	KNN	0.011	0.000	0.000	0.000
	RF	1.000	0.000	0.245	0.007
	SVM	0.010	0.000	0.000	0.000
	ANN	0.000	0.000	0.001	0.000
	Multinorm	0.000	0.000	0.000	0.000
	LDA	0.000	0.000	0.000	0.000
	Ordinal	0.000	0.000	0.000	0.000
	CART	0.000	0.000	0.000	0.000
KNN	RF	0.003	0.281	0.000	0.041
	SVM	1.000	1.000	0.000	1.000
	ANN	0.231	0.000	1.000	0.000
	Multinorm	0.000	0.000	0.779	0.000
	LDA	0.000	0.000	1.000	0.000
	Ordinal	0.000	0.000	0.000	0.000
	CART	0.000	0.000	0.004	0.000
RF	SVM	0.203	0.014	0.008	0.002
	ANN	0.002	0.000	0.002	0.000
	Multinorm	0.000	0.000	0.000	0.000
	LDA	0.000	0.000	0.000	0.000
	Ordinal	0.000	0.000	0.000	0.000
	CART	0.000	0.000	0.000	0.000
SVM	ANN	0.563	0.000	0.021	0.000
	Multinorm	0.000	0.000	0.000	0.000
	LDA	0.000	0.000	0.000	0.000
	Ordinal	0.000	0.000	0.000	0.000
	CART	0.000	0.000	0.000	0.000
ANN	Multinorm	0.002	0.000	1.000	0.000
	LDA	0.000	0.000	1.000	0.000
	Ordinal	0.002	0.000	0.000	0.000
	CART	0.000	0.000	0.903	0.000
Multinorm	LDA	0.019	1.000	1.000	1.000
	Ordinal	0.004	0.000	0.000	0.000
	CART	0.000	0.000	0.023	0.000
LDA	Ordinal	0.019	0.000	1.000	0.006
	CART	0.000	0.000	0.032	0.000
Ordinal	CART	0.000	0.002	0.047	0.008

Model: (XGB: Gradient boosting decision tree model, RF: Random Forest, KNN: K-Nearest Neighbours, SVM: Support Vector Machines, ANN: Neural Networks, Multinorm: Multinomial regression, LDA: Linear Discriminant Analysis, Ordinal: ordinal logistic regression, CART: Classification and regression tree).



Appendix IX Figure 2 Random Forest based Multi-dimensional score (MDS) plots for BCS prediction in 43−54 months old ewes at different stages of the annual cycle (a: pre-breeding, b: pregnancy diagnosis, c: pre-lambing, d: weaning). Red, blue and green, circles represent single data points from BCS of 1.0−2.0, 2.5−3.5 and > 3.5, respectively.

Appendix X: Frequency of ewe BCS, correlation between BCS, LW and HW, and BCS prediction error (Chapter 9)

Appendix X Table 1 Frequency of ewe body condition scores (BCS) by age (8−18 months, 19−30, 31−42) and stage of the annual production cycle.

	8−18				19−30				31−42			
BCS	PB	PD	PL	W	PB	PD	PL	W	PD	PL	W	Overall
1.5				5	6	1	2	5	2		17	5
2.0	34	38	13	47	32	5	8	84	16	7	67	32
2.5	186	208	180	175	111	65	89	160	114	68	140	136
3.0	156	145	171	135	131	136	127	97	119	87	63	124
3.5	45	36	53	43	104	111	106	41	83	119	45	71
4.0	6	2	11	22	32	67	65	20	49	76	14	33
4.5				2	11	37	24	16	19	36	16	20
5.0				1		4	3	3		5	10	4

Stage of the annual production cycle (PB: pre-breeding, PD: pregnancy diagnosis, PL: pre-lambing, W: weaning). Empty space indicates no ewe had that body condition score.

Appendix X Table 2 Correlation coefficients between individual unadjusted and/or adjusted liveweight (LW) and height at withers (HW) across stages of the annual production cycle in ewes between 8 and 42 months.

Wither Height measurement	n	8–18 PB	8–18 PD	8–18 PL	8–18 W	19–30 PB	19–30 PD	19–30 PL	19–30 W	31–42 PB	31–42 PD	31–42 PL	31–42 W
Unadjusted LW													
LH[1]	428	0.57	0.56	0.57	0.47	-0.24	0.44	0.45	0.34	0.32	0.44	0.43	0.29
PH[2]	427	0.27	0.23	0.19	0.30	-0.15	0.28	0.29	0.17	0.14	0.20	0.22	0.19
DH[2]	426	0.37	0.35	0.33	0.38	-0.14	0.43	0.39	0.26	0.24	0.32	0.34	0.26
WH[2]	424	0.37	0.31	0.29	0.34	-0.17	0.30	0.26	0.41	0.36	0.41	0.37	0.32
DH[3]	402	0.30	0.25	0.20	0.24	-0.18	0.24	0.20	0.25	0.20	0.29	0.27	0.19
WH[3]	402	0.35	0.30	0.30	0.30	-0.12	0.34	0.30	0.36	0.33	0.38	0.34	0.35
Adjusted LW													
LH[1]	428	0.57	0.56	0.57	0.47	-0.25	0.45	0.48	0.34	0.32	0.44	0.41	0.29
PH[2]	427	0.27	0.24	0.19	0.30	-0.16	0.27	0.28	0.16	0.13	0.19	0.23	0.19
DH[2]	426	0.37	0.35	0.33	0.38	-0.15	0.43	0.40	0.26	0.23	0.31	0.33	0.26
WH[2]	424	0.37	0.32	0.29	0.34	-0.17	0.34	0.38	0.41	0.36	0.41	0.38	0.32
DH[3]	402	0.30	0.26	0.20	0.23	-0.19	0.25	0.25	0.24	0.19	0.28	0.26	0.18
WH[3]	402	0.34	0.30	0.29	0.29	-0.12	0.37	0.39	0.36	0.32	0.37	0.38	0.34

Adjusted indicates that variables were corrected for fleece conceptus and fleece weight. All correlation coefficients were significant ($p < 0.05$). PB, PD, PL, W indicate the four stages of the annual production cycle including pre-breeding, pregnancy diagnosis, pre-lambing, and weaning, respectively. PH: height at withers at breeding, DH: pregnancy diagnosis, LH: pre-lambing, and WH: weaning. Correlations based on centred and scaled training data. Underlined values indicate correlation of both HW and LW measurement at the same time point. [1,2,3] denote the time (Months: 8-18, 19-30, 31-42, respectively) of measurement.

Appendix X Table 3 Correlation coefficients between individual unadjusted or adjusted liveweight (LW) and body condition scores (BCS) across stages of the annual production cycle in ewes between 8 and 42 months.

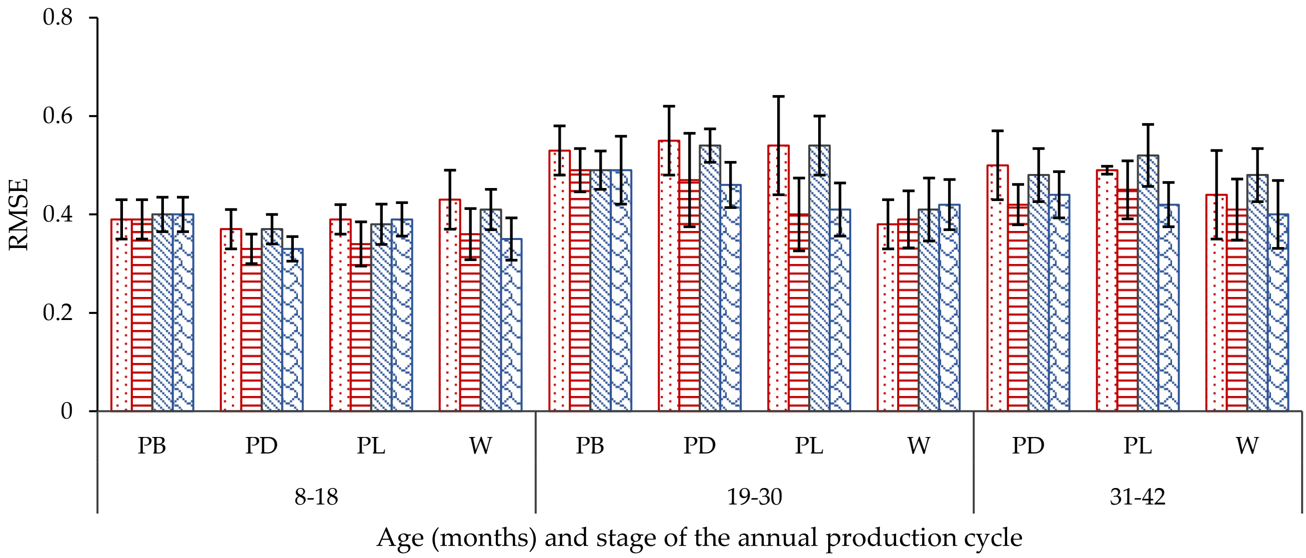
		8–18				19–30				31–42		
predictor	n	PB	PD	PL	W	PB	PD	PL	W	PD	PL	W
						Liveweight alone1 (Unadjusted)						
WM1	428	0.40**	0.26**	0.08	0.06	-0.11*	0.08	0.11*	0.01	0.10*	0.13*	0.05
WP1	429	0.38**	0.34**	0.11*	0.07	-0.08	0.11*	0.16**	0.01	0.12*	0.18**	0.06
WL1	428	0.36**	0.28**	-0.01	-0.03	-0.12*	0.05	0.12*	0.05	0.16**	0.20**	0.06
WW1	429	0.19**	0.19**	0.15**	0.41**	-0.03	0.35**	0.33**	0.07	0.23**	0.28**	0.20**
WM2	427	0.14**	-0.05	0.03	0.16**	0.54**	0.11*	0.07	-0.02	0.01	-0.03	0.10
WP2	426	0.16**	0.20**	0.21**	0.36**	0.01	0.45**	0.34**	-0.06	0.18**	0.21**	0.18**
WL2	424	0.15**	0.19**	0.17**	0.34**	0.01	0.40**	0.26**	-0.16**	0.06	0.15**	0.12*
WW2	424	0.02	0.02	0.06	0.14**	-0.06	0.22**	0.39**	0.68**	0.55**	0.49**	0.33**
WP3	402	0.06	0.08	0.04	0.13*	-0.06	0.23**	0.35**	0.44**	0.55**	0.53**	0.28**
WL3	399	0.06	0.07	0.03	0.10	-0.05	0.21**	0.30**	0.29**	0.43**	0.45**	0.16**
WW3	402	0.05	0.06	0.04	0.18**	-0.03	0.27**	0.28**	0.31**	0.41**	0.45**	0.72**
						Liveweight alone2 (Adjusted)						
WM[1]	428	0.40**	0.25**	0.08	0.07	-0.11*	0.08	0.12*	0.02	0.12*	0.14**	0.05
WP[1]	429	0.36**	0.34**	0.14**	0.10	-0.08	0.15**	0.19**	0.02	0.14**	0.20**	0.07
WL[1]	428	0.37**	0.28**	-0.01	-0.03	-0.11*	0.05	0.12*	0.05	0.16**	0.20**	0.06
WW[1]	429	0.19**	0.18**	0.14**	0.41**	-0.03	0.35**	0.33**	0.08	0.23**	0.28**	0.20**
WM[2]	427	0.14**	-0.05	0.03	0.15**	0.53**	0.10*	0.07	-0.02	0.01	-0.02	0.10
WP[2]	426	0.15**	0.20**	0.22**	0.37**	0.02	0.47**	0.40**	-0.03	0.25**	0.27**	0.21**
WL[2]	424	0.12*	0.19**	0.19**	0.35**	0.02	0.44**	0.44**	-0.10*	0.25**	0.31**	0.22**
WW[2]	424	0.02	0.02	0.06	0.14**	-0.05	0.22**	0.39**	0.68**	0.55**	0.49**	0.32**
WP[3]	402	0.06	0.08	0.05	0.13*	-0.05	0.23**	0.36**	0.46**	0.56**	0.55**	0.31**
WL[3]	399	0.05	0.09	0.05	0.12*	-0.03	0.24**	0.34**	0.33**	0.46**	0.53**	0.29**
WW[3]	402	0.05	0.07	0.04	0.18**	-0.03	0.27**	0.28**	0.31**	0.41**	0.45**	0.72**

Adjusted indicates that variables were corrected for fleece conceptus and fleece weight. Asterisks *, ** indicate significance at $p<0.05$ and $p<0.01$ respectively. PB, PD, PL, W indicate the four stages of the annual production cycle including pre-breeding, pregnancy diagnosis, pre-lambing, and weaning, respectively. Underlined values indicate correlation of both BCS and LW measurement at the same time point. Correlations based on centred and scaled training data. [1,2,3] denote the time (Months: 8-18, 19-30, 31-42, respectively) of measurement.

Appendix X Table 4 Correlation coefficients between individual height at withers (HW) and body condition scores (BCS) across stages of the annual production cycle in ewes between 8 and 42 months.

		8–18				19–30				31–42		
HW	n	PB	PD	PL	W	PB	PD	PL	W	PD	PL	W
						Unadjusted WH						
LH[1]	428	0.08	0.11*	-0.06	-0.04	-0.02	-0.02	-0.02	0.05	0.08	0.04	0.00
PH[2]	427	-0.03	0.02	0.00	0.08	-0.06	-0.02	-0.03	-0.05	0.02	0.01	0.07
DH[2]	426	0.09	0.10*	-0.05	0.01	-0.02	0.07	0.06	-0.01	0.05	0.02	0.00
WH[2]	424	-0.03	0.01	-0.06	0.00	-0.06	-0.04	0.06	0.15**	0.09	0.05	0.09
DH[3]	402	0.02	-0.01	-0.01	0.04	-0.06	-0.01	-0.03	0.07	0.07	0.13*	0.03
WH[3]	402	0.03	0.04	-0.02	-0.01	-0.02	0.02	0.04	0.06	0.12*	0.14**	0.14**
						Adjusted HW						
LH[1]	428	0.08	0.11*	-0.06	-0.04	-0.02	-0.02	-0.02	0.05	0.08	0.04	0.00
PH[2]	427	-0.03	0.02	0.00	0.08	-0.06	-0.02	-0.03	-0.05	0.02	0.01	0.07
DH[2]	426	0.09	0.10*	-0.05	0.01	-0.02	0.07	0.06	-0.01	0.05	0.02	0.00
WH[2]	424	-0.03	0.01	-0.06	0.00	-0.06	-0.04	0.06	0.15**	0.09	0.05	0.09
DH[3]	402	0.02	-0.01	-0.01	0.04	-0.06	-0.01	-0.03	0.07	0.07	0.13*	0.03
WH[3]	402	0.03	0.04	-0.02	-0.01	-0.02	0.02	0.04	0.06	0.12*	0.14**	0.14**

Adjusted indicates that variables were corrected for fleeceweight. Asterisks *, ** indicate significance at p<0.05 and p<0.01 respectively. PB, PD, PL, W indicate the four stages of the annual production cycle including pre-breeding, pregnancy diagnosis, pre-lambing, and weaning, respectively. PH: height at withers at breeding, DH: pregnancy diagnosis, LH: pre-lambing, and WH: weaning. Underlined values: same time point correlation coefficients. Correlations based on centred and scaled training data. [1,2,3] denote the time (Months: 8-18, 19-30, 31-42, respectively) of measurement.



Appendix X Figure 1 Root mean square error (RMSE with standard deviations) of models (dotted bar: unadjusted liveweight alone models, horizontal stripes: combined models based on unadjusted LW, liveweight change and previous BCS, diagonal stripes: adjusted liveweight alone, shingled: adjusted liveweight, liveweight change, height at withers and previous BCS) for current BCS prediction across the stage of the annual production cycle and ewe age group. Colours (Red indicates unadjusted liveweight while blue indicates adjusted liveweight was used). PB, PD, PL, W indicate body condition score prior to pre-breeding, at pregnancy diagnosis, prior to lambing, and at weaning, respectively.

DRC 16

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Jimmy Semakula
Name/title of Primary Supervisor:	Professor Paul R Kenyon
In which chapter is the manuscript /published work:	Chapter 3

Please select one of the following three options:

(●) The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

Semakula, J., Corner-Thomas, R., Morris, S., Blair, H., and Kenyon, P. (2019). The effect of herbage type prior to fasting on the rate of liveweight loss during fasting in ewe lambs. New Zealand Journal of Animal Science and Production 79, 131-134.

( ) The manuscript is currently under review for publication – please indicate:

- The name of the journal:
- The percentage of the manuscript/published work that was contributed by the candidate:
- Describe the contribution that the candidate has made to the manuscript/published work:

( ) It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Jimmy Semakula Digitally signed by Jimmy Semakula Date: 2021.03.01 21:10:42 +13'00'
Date:	01-Mar-2021
Primary Supervisor's Signature:	Paul Kenyon Digitally signed by Paul Kenyon DN: cn=Paul Kenyon, c=NZ, o=Massey University, ou=SAE, email=p.r.kenyon@massey.ac.nz Date: 2021.03.05 11:53:18 +13'00'
Date:	

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

GRS Version 5 – 13 December 2019
DRC 19/09/10

DRC 16

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Jimmy Semakula
Name/title of Primary Supervisor:	Professor Paul R Kenyon
In which chapter is the manuscript /published work:	Chapter 4

Please select one of the following three options:

(●) The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

Semakula, J., Corner-Thomas, R. A., Morris, S. T., Blair, H. T., and Kenyon, P. R. (2021). The Effect of Herbage Availability and Season of Year on the Rate of Liveweight Loss during Weighing of Fasting Ewe Lambs. Agriculture 11, 150.

( ) The manuscript is currently under review for publication – please indicate:

- The name of the journal:
- The percentage of the manuscript/published work that was contributed by the candidate:
- Describe the contribution that the candidate has made to the manuscript/published work:

( ) It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Jimmy Semakula Digitally signed by Jimmy Semakula Date: 2021.03.01 21:10:42 +13'00'
Date:	01-Mar-2021
Primary Supervisor's Signature:	Paul Kenyon Digitally signed by Paul Kenyon DN: cn=Paul Kenyon, c=NZ, o=Massey University, ou=SAE, email=p.r.kenyon@massey.ac.nz Date: 2021.03.05 11:53:58 +13'00'
Date:	

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

GRS Version 5 – 13 December 2019
DRC 19/09/10

DRC 16

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Jimmy Semakula
Name/title of Primary Supervisor:	Prof. Paul R Kenyon
In which chapter is the manuscript /published work:	Chapter 5

Please select one of the following three options:

(●) The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

Semakula J, Corner- Thomas R.A, R, Morris S.T, Blair, H.T, Kenyon, P.R. The Effect of Herbage Availability, Pregnancy Stage and Rank on the Rate of Liveweight Loss during Fasting in Ewes. Agriculture 2021, 11, 543.

( ) The manuscript is currently under review for publication – please indicate:

- The name of the journal:
- The percentage of the manuscript/published work that was contributed by the candidate:
- Describe the contribution that the candidate has made to the manuscript/published work:

( ) It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Jimmy Semakula Digitally signed by Jimmy Semakula Date: 2021.07.07 16:32:22 +03'00'
Date:	
Primary Supervisor's Signature:	PRKyn
Date:	8/7/2021

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

GRS Version 5 – 13 December 2019
DRC 19/09/10

DRC 16

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Jimmy Semakula
Name/title of Primary Supervisor:	Professor Paul R Kenyon

In which chapter is the manuscript /published work: Chapter 6

Please select one of the following three options:

(●) The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

Semakula, J., Corner-Thomas, R. A., Morris, S. T., Blair, H. T., and Kenyon, P. R. (2020a). The Effect of Age, Stage of the Annual Production Cycle and Pregnancy-Rank on the Relationship between Liveweight and Body Condition Score in Extensively Managed Romney Ewes. Animals 10, 784.

( ) The manuscript is currently under review for publication – please indicate:

- The name of the journal:
- The percentage of the manuscript/published work that was contributed by the candidate:
- Describe the contribution that the candidate has made to the manuscript/published work:

( ) It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Jimmy Semakula Digitally signed by Jimmy Semakula Date: 2021.03.01 21:10:42 +13'00'
Date:	01-Mar-2021
Primary Supervisor's Signature:	Paul Kenyon Digitally signed by Paul Kenyon DN: cn=Paul Kenyon, c=NZ, o=Massey University, ou=SAE, email=p.r.kenyon@massey.ac.nz Date: 2021.03.03 11:54:13 +13'00'
Date:	

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

GRS Version 5 – 13 December 2019
DRC 19/09/10

DRC 16

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Jimmy Semakula
Name/title of Primary Supervisor:	Professor Paul R Kenyon
In which chapter is the manuscript /published work:	Chapter 7

Please select one of the following three options:

(●) The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

Semakula, J., Corner-Thomas, R. A., Morris, S. T., Blair, H. T., and Kenyon, P. R. (2020b). Predicting Ewe Body Condition Score Using Lifetime Liveweight and Liveweight Change, and Previous Body Condition Score Record. Animals 10, 1182.

( ) The manuscript is currently under review for publication – please indicate:

- The name of the journal:
- The percentage of the manuscript/published work that was contributed by the candidate:
- Describe the contribution that the candidate has made to the manuscript/published work:

( ) It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Jimmy Semakula Digitally signed by Jimmy Semakula Date: 2021.03.01 21:10:42 +13'00'
Date:	01-Mar-2021
Primary Supervisor's Signature:	Paul Kenyon Digitally signed by Paul Kenyon DN: cn=Paul Kenyon, c=NZ, o=Massey University, ou=SAE, email=p.r.kenyon@massey.ac.nz Date: 2021.03.05 11:54:24 +13'00'
Date:	

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

GRS Version 5 – 13 December 2019
DRC 19/09/10

DRC 16

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Jimmy Semakula
Name/title of Primary Supervisor:	Professor Paul R Kenyon
In which chapter is the manuscript /published work:	Chapter 8

Please select one of the following three options:

(●) The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

Semakula J, Corner-Thomas RA, Morris ST, Blair HT, Kenyon PR. Application of Machine Learning Algorithms to Predict Body Condition Score from Liveweight Records of Mature Romney Ewes. Agriculture. 2021; 11(2):162.

( ) The manuscript is currently under review for publication – please indicate:

- The name of the journal:
- The percentage of the manuscript/published work that was contributed by the candidate:
- Describe the contribution that the candidate has made to the manuscript/published work:

( ) It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Jimmy Semakula Digitally signed by Jimmy Semakula Date: 2021.03.01 21:10:42 +13'00'
Date:	01-Mar-2021
Primary Supervisor's Signature:	Paul Kenyon Digitally signed by Paul Kenyon DN: cn=Paul Kenyon, c=NZ, o=Massey University, ou=SAE, email=p.r.kenyon@massey.ac.nz Date: 2021.03.05 11:54:37 +13'00'
Date:	

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

GRS Version 5 – 13 December 2019
DRC 19/09/10

DRC 16

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Jimmy Semakula
Name/title of Primary Supervisor:	Prof. Paul R Kenyon
In which chapter is the manuscript /published work:	Chapter 9

Please select one of the following three options:

○ The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

◉ The manuscript is currently under review for publication – please indicate:

- The name of the journal:

Translational Animal Science

- The percentage of the manuscript/published work that was contributed by the candidate: 60.00

- Describe the contribution that the candidate has made to the manuscript/published work:

Conceptualization, Data collection, curation, software, formal analysis, results interpretation, validation and preparation of the manuscript.

○ It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Jimmy Semakula Digitally signed by Jimmy Semakula Date: 2021.07.07 16:40:32 +03'00'
Date:	
Primary Supervisor's Signature:	PRKenyon
Date:	8/7/2021

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

GRS Version 5 – 13 December 2019
DRC 19/09/10