

Application of random regression models to study growth curves and genotype x environment interactions in sheep.

Emma McGowan

A thesis submitted for the degree of Doctor of Philosophy

College of Medicine and Veterinary Medicine

Global Academy of Agriculture and Food Systems

University of Edinburgh

September 2024



## Declaration

I declare that this thesis is my own composition and that the research described in it is my own work, except where otherwise stated.

Emma McGowan Date: 30 September 2024

## ABSTRACT

Modelling sheep growth and producing estimated breeding values (EBV) for growth traits are widely used to optimize sheep production. The methods available model growth traits as a function of age, often with a set of fixed or random environmental effects. Current methods to map growth rely on two prominent assumptions that: (1) The mean and covariance structure of the growth trait remains constant with time or age, and (2) each measurement of the growth trait is genetically different from, but correlated to, all other measurements of the growth trait. However, this methodology is problematic. It has three main issues: (1) impracticality; (2) neglecting wide random variability in environmental effects; and (3) overparameterization. Thus, this project used a random regression model to describe growth more accurately in sheep, thereby producing better selection criteria to choose the best breeding stock. These models included genetic effects, maternal genetic effects, and a host of conformation and meat quality data. However, random regression models can struggle to provide accurate results when data are sparse or unevenly distributed. Thus, the use of commercial data in this project allows investigation of model performance and assessment of the applicability of random regression models in commercial environments. Overall, the project found random regression models are highly sensitive to the distribution and variability of records across the age distribution. This results in high correlations between the parameters of the regression model which can result in inaccurate genetic parameters. In Chapter 2 random regression models were constructed for growth in Charollais and Suffolk sheep with constrained correlations resulting in heritabilities between (0.18-0.49) and (0.20-0.50) respectively. The inclusion of the constrained correlation was validated using a novel procedure. In Chapter 3, carcass information was incorporated using commercial mixed breed and Scottish Blackface research data with similar numbers of records. A selection index and a random regression model were compared. Model convergence for the random regression model was achieved by constraining the correlation. Genetic parameters could be calculated between weights and fat class or conformation. The selection index offered accurate information for slaughter weight and carcass weight. Chapter 4 assessed GxE effects in the RamCompare project using a sire model with a sire by flock interaction and a reaction norm model using a phenotypic deviation. It indicated GxE effects were present for birth weight, scan weight, weaning weight and muscle depth. It also showed that an assessment of GxE effects in commercial flocks can be conducted in datasets that lack environmental data. The project contributes to a growing body of research on how to best model heritable traits and provide genetic information to commercial sheep flocks.

## LAY SUMMARY

Sheep are a large component of the UK's agricultural output and food system. However, lamb production is challenged by increasing global demand for animal protein, emissions reduction and demands for increased productivity. One of the major ways that animal science addresses these challenges is by focusing on optimizing animal growth. Animal growth is an optimal target for research because it can be influenced by an animal's genetics. Contemporary genetics can identify animals that have the potential to grow efficiently and to pass on this trait to their offspring. Genetic modelling is one of the ways to identify the animals that have the potential to grow quickly in many different environments. The standard methodology models weight as a function of age and include environmental variables to account for the influence the environment may have on growth. This methodology relies on two key assumptions. The first is that the average growth rate remains constant from the beginning of the growth cycle to the end, and that each measurement of growth is a distinct trait that is related to the other measurements. Models with multiple weight measurements are used to estimate values that represent the animal's genetic potential for growth compared to an average value for the population. However, these models have some major drawbacks. They do not account for the changeability in growth over an animal's entire life. They can only account for the discrete moments in time where a weight measurement is taken and the environmental effects that are at play in that single moment. This project uses random regression models to more accurately model the entire growth cycle. The project also investigates some of the areas where random regression models struggle to function well. These include areas where there are not many records or where those weight records are not evenly distributed throughout an animal's life. This project not only estimates genetic values for growth throughout an animal's entire life but also seeks to investigate solutions to the pitfalls of random regression modelling.

## CONTENTS

Abstract .....	i
Lay Summary .....	ii
Contents.....	iii
Table of Contents .....	iv
List of Figures .....	vii
List of Tables.....	viii
Acknowledgements.....	x
Dedication.....	xi

## TABLE OF CONTENTS

<b>Chapter One: Introduction and Background .....</b>	<b>Page 1</b>
1.1 Introduction to the UK Sheep Industry Introduction to the UK Sheep Industry.....	Page 1
1.2 Growth Models in UK Sheep Production.....	Page 4
1.3 Genetic Modelling and Selection of UK Sheep.....	Page 5
1.4 Fixed Point Growth Modelling.....	Page 9
1.5 Random Regression Model Definitions and Applications.....	Page 10
1.5.1 Definition of the Random Regression Model .....	Page 10
1.5.2 Application of Random Regression Models.....	Page 13
1.6 The Aims of This Thesis and Contribution to the Producer.....	Page 17
 <b>Chapter Two: Modelling growth in Suffolk and Charollais sheep populations with insufficient mature weight records for polynomial random regression model.....</b>	 <b>Page 20</b>
2.1 Introduction.....	Page 20
2.2 Published Paper Copy.....	Page 22
2.3 Discussion.....	Page 31
 <b>Chapter Three: Using Random Regression Models to Model Growth and Environmental Effect in Sheep with Irregular Data.....</b>	 <b>Page 33</b>
3.1 Introduction.....	Page 33
3.1.1 Random Regression Modelling with Unevenly Distributed Data.....	Page 33
3.1.2 Including Carcass Traits in Genetic Models .....	Page 34
3.2 Materials and Methods.....	Page 35
3.2.1 Phenotypic Curves.....	Page 40
3.2.2 Genetic Analysis.....	Page 40
3.2.3 Relationship Between Growth and Carcass Traits for Scottish Blackface Sheep .....	Page 44
3.2.3.1 Selection Index.....	Page 44
3.2.3.2 Random Regression Analysis of Carcass Traits in SBF.....	Page 47
3.3 Results.....	Page 48
3.3.1 Phenotypic Curves.....	Page 48
3.3.2 Genetic Analysis.....	Page 49
3.3.3 Relationship between Growth and Carcass Traits.....	Page 51
3.3.3.1 Selection Index.....	Page 51
3.3.3.2 Random Regression Modelling of Growth and Carcass Traits	Page 55

3.4 Discussion.....	Page 59
3.4.1 Phenotypic Curves.....	Page 59
3.4.2 Genetic Analysis.....	Page 59
3.4.3 Relationship between Growth and Carcass Traits.....	Page 61
3.4.3.1 Selection Index.....	Page 61
3.4.3.2 Random Regression Analysis of Carcass Traits in SBF.....	Page 67

<b>Chapter Four: Modelling GxE interaction in terminal sire breeds in UK sheep populations.....</b>	<b>Page 72</b>
4.1 Introduction.....	Page 72
4.2 Materials and Methods.....	Page 77
4.2.1 Data Collection and Experimental Design.....	Page 77
4.2.2 Sire and Flock Interaction.....	Page 77
4.2.3 Reaction Norm Modelling.....	Page 79
4.3 Results.....	Page 80
4.3.1 Sire and Flock Interaction.....	Page 80
4.3.2 Reaction Norm Modelling.....	Page 85
4.4 Discussion.....	Page 90
4.4.1 Sire and Flock Interaction.....	Page 92
4.4.2 Reaction Norm Modelling.....	Page 94
4.4.3 Implications for RamCompare Breeders.....	Page 96

<b>Chapter Five: General Discussion.....</b>	<b>Page 98</b>
5.1 Using RRM for Growth in Sheep.....	Page 99
5.2 Relationship between Growth and Non-longitudinal Traits.....	Page 102
5.3 Optimizing Selection Criteria for Growth and Carcass Traits.....	Page 104
5.4 Estimate GxE interaction in UK Sheep Breeds.....	Page 107

## LIST OF FIGURES

- Figure 3.1:** The EUROP Grid used for sheep carcass grading.....Page 34
- Figure 4.1** Genotype response with GxE causing scaling but no reranking of  
Genotypes.....Page 72
- Figure 4.2:** Genotype response with GxE causing reranking of genotypes..... Page 73
- Figure 4.3** Ranked EBVs for Muscle Depth for the same sires between two flocks showing  
reranking.....Page 81
- Figure 4.4** Ranked EBVs for Muscle Depth for the same sires between two flocks showing  
minimal or no reranking .....Page 82
- Figure 4.5** Ranked EBVs for Scan Weight for the same sires between two flocks showing  
reranking.....Page 82
- Figure 4.6** Ranked EBVs for Scan Weight for the same sires between two flocks showing  
minimal or no reranking..... Page 83
- Figure 4.7** Ranked EBVs for Weaning Weight for the same sires between two flocks showing  
reranking..... Page 83
- Figure 4.8** Ranked EBVs for Weaning Weight for the same sires between two flocks  
showing no or minimal reranking .....Page 84
- Figure 4.9** Ranked EBVs for Birth Weight for the same sires between two flocks showing  
reranking.....Page 84
- Figure 4.10** Ranked EBVs for Birth Weight for the same sires between two flocks showing  
no or minimal reranking. ....Page 85
- Figure 4.11** The phenotypic average for Muscle Depth for the top three sires across  
flocks.....Page 88
- Figure 4.12** The phenotypic average for Weaning Weight for the top three sires  
across flocks.....Page 89
- Figure 4.13** The phenotypic average for Scan Weight for the top three sires across  
flocks.....Page 89

## LIST OF TABLES

<b>Table 1:</b> Summary of records for each weight recording point in the Suffolk and Charollais datasets .....	Page 24
<b>Table 2:</b> Summary of variables for Suffolk and Charollais datasets .....	Page 24
<b>Table 3:</b> Genetic (below diagonal) and phenotypic (above diagonal) correlations among several ages in days (along the top and left side) with heritabilities on the diagonal from the random regression model for Method B for Suffolk .....	Page 26
<b>Table 4:</b> Genetic (below diagonal) and phenotypic (above diagonal) correlations among several ages in days (along the top and left side) with heritabilities on the diagonal from the random regression model for Method B for Charollais. ....	Page 27
<b>Table 5:</b> Validation results for fixed correlation between intercept and slope at three different levels for the Suffolk breed. ....	Page 27
<b>Table 6:</b> Validation results for fixed correlations between intercept and slope at three different levels for the Charollais breed. ....	Page 28
<b>Table 3.1:</b> Display of variables in the RC dataset. ....	Page 37
<b>Table 3.2:</b> A description of the variables in SBF. ....	Page 38
<b>Table 3.3</b> A summary of carcass traits for the SBF dataset.....	Page 46
<b>Table 3.4:</b> Phenotypic model summary values.....	Page 50
<b>Table 3.5:</b> Genetic (below diagonal) and Phenotypic (above diagonal) correlations among several weights at different ages with heritabilities (bold) on the diagonal from the random regression model for RC. ....	Page 50
<b>Table 3.6:</b> Genetic (below diagonal) and Phenotypic (above diagonal) correlations among several ages in days with heritabilities (bold) on the diagonal from the random regression model for SBF Data. ....	Page 51
<b>Table 3.7:</b> G Matrix calculated for the Scottish Blackface data from multivariate models.....	Page 51
<b>Table 3.8:</b> The R matrix calculated for the Scottish Blackface data from multivariate models.....	Page 52
<b>Table 3.9:</b> The P matrix calculate for the Scottish Blackface data from multivariate models.....	Page 52
<b>Table 3.10:</b> Selection index results for slaughter weight. ....	Page 53

**Table 3.11:** The selection index results for carcass weight.....Page 54

**Table 3.12:** The results for the selection index for conformation .....Page 54

**Table 3.13:** The results for the selection index for fat class.....Page 55

**Table 3.14:** A table of heritabilities down the diagonal with genetic correlations below the diagonal and phenotypic correlations above the diagonal for weights throughout an age distribution of 1 to 376 and Conformation .....Page 57

**Table 3.15:** A table of heritabilities down the diagonal with genetic correlations below the diagonal and phenotypic correlations above the diagonal for weights throughout an age distribution of 1 to 376 and Fat Class.....Page 58

**Table 4.1** Heritability values across flocks for weaning weight .....Page 86

**Table 4.2** Heritability values across flocks for scan weight .....Page 86

**Table 4.3** Heritability values across flocks for muscle depth.....Page 86

**Table 4.4** Top three ranked Sires and EBVs from Reaction Norm Analysis of Muscle Depth.....Page 87

**Table 4.5** Top three ranked Sires and EBVs from Reaction Norm Analysis of Weaning Weight.....Page 87

**Table 4.6** Top three ranked Sires and EBVs from Reaction Norm Analysis of Scan Weight.....Page 87

## ACKNOWLEDGEMENTS

Thank you first to AHDB for funding the PhD studentship and providing a great deal of the data for analysis. A special thank you to Kim Matthews and Samuel Boon at AHDB who supported the project but also connected me with great researchers near and far.

I am extremely grateful to my supervisors Raphael Mrode, Mike Coffey, and Geoff Simm. They have been patient and supportive as I learned new skills. They have encouraged me as I brought the research to conferences and publications. My supervisors were also incredibly helpful navigating my PhD while coming back from maternity leave. Thank you for your patience and understanding.

During my research, a number of people offered their helpful advice and support. Thank you to Harriet Bunning for her help preparing for presentations. Thank you to Nicola Lambe and Joanne Conington who supplied data and advised on analysis of the Scottish Blackface datasets. Thank you to Ann McLaren who supplied data for the Scottish Blackface random regression modelling with variable numbers of weight recordings. Thank you all for your personal encouragement as well as help with data and analysis.

Finally, thank you to my family, Mom, Dad, Grace, Jamie and Seumas. Your love and support have meant the world to me throughout this process. I hope this makes you proud! Seumas, you made all this more meaningful than it could have ever been without you. Jamie, thank you for sitting up nights and waking early, for endless coffees and pep talks, thank you for believing in me.

**DEDICATION**

**For my Mom,**

**Thank you for everything,**

**Especially for letting me run the Quad J Ranch Sheep Rescue.**



## **CHAPTER ONE: INTRODUCTION**

### **1.1 INTRODUCTION TO THE UK SHEEP INDUSTRY**

The UK sheep industry includes approximately 33 million sheep in about 26,287 businesses with about 1.8 billion pounds in revenue for the year 2024 (AHDB, 2021). This sizeable portion of the agriculture sector is dominated by rearing lambs for slaughter to produce meat over other parts of the sector like wool production (IBIS World, 2024). In terms of breed profiles and genetic makeup, the UK sheep sector has been traditionally categorized as a “stratified crossbreeding” population. This means that terminal sire breeds are crossed with a fair variety of lowland purebred ewes (AHDB, 2021). This fairly unique structure allows for sheep farming to maintain in relatively inhospitable environments while producing replacement hill ewes and ram lambs, lightweight store lambs and cast ewes. For upland farming, ewes are crossed with long wool rams to produce hybrid ewe lambs and finishing wether lambs. Hybrid ewes with desirable maternal traits like Mules and Greyfaces are bred with terminal sire rams like Beltex, Charollais, Dorset, or Texel etc. to produce most UK finishing lambs. As of 2021, terminal sires dominated the genetic profile of the UK sheep population with 58% of the lamb crop being sired by terminal breeds. The top three most popular sire breeds in the UK are Texel, Suffolk and Charollais as of 2021 (AHDB, 2021).

The main advantage of the stratified system is that it exploits the best characteristics of the wide range of breeds in the UK. In 2021 alone, 99 different breeds were recorded in the UK sheep population (AHDB, 2021). The best characteristics of these breeds are incorporated into the genetic profile of the UK sheep population through the stratified breeding system. Greyface and Mule ewes combine long wool traits with good growth rates and manageable litter size.

The hardiness of the hill breeds and manageable mature size enables the use of terminal sires to produce a lamb that is tailored to the climate and to the market. The main drawback is the number of movements this stratified breeding system requires which can increase the risk of disease. Additionally, most lamb finishers rely on the replacement of crossbred ewe lambs making it difficult to select for certain traits like milk production, growth rate, and disease resistance (Sargison, 2009; AHDB 2021).

The breeding system nonetheless results in relatively easily managed and quickly growing finishing lambs. They are sent to slaughter at an average of 6 to 7 months of age. The average slaughter weight is around 40 to 45 kg depending on the age and breed of the lamb (AHDB 2021). The UK lamb sector is highly motivated to achieve this slaughter weight as economically and efficiently as possible because this is the point of income for the producer. The current genetic evaluation for sheep in the UK also aims to achieve this outcome by highlighting both terminal traits in the terminal index, focusing on live weights and carcass traits, and maternal traits in the maternal index, focusing on maternal ability and litter sizes, which contribute to lambs that can achieve profitable outcomes. The sheep evaluation therefore includes birth weight, lambing ease, lamb survival, 8 week weight, shearling weight, litter size, litter size reared, maternal ability, scan weight, muscle depth, fat depth, days to slaughter, carcass weight, carcass conformation, and fat class (Signet, 2021).

Sheep growth is therefore an important trait both for economic return and resource use efficiency. Accurately measuring and predicting growth curves is therefore an important way to understand how quickly and efficiently animals grow in relation to available nutrients both in timing and nutrient density. Animal growth is not only an important trait for geneticists or farmers, but it has global significance to the development of society and the welfare of

humanity. Animal products are part of the food supply chain across the world making their production a core component or providing safe, traceable, nutritious food for humans across the world. Animal production is also under increasing pressure to deliver other public goods by minimizing environmental impact of animal production and using techniques that enhance or rebuild the environments in which animals are reared. There are many ways to do this including increasing the efficiency of animal production. Selection for animal growth is a key part of improving overall system efficiency. For example, animals that reach slaughter weight earlier emit less greenhouse gas during their life. As of 2019, animal products from primary species such as cattle, chicken, and sheep made up 40% of global agricultural output, 18% of global calorie intake, and 40% of global protein intake (ILRI, 2019).

Thus, animal science has an important role to play in understanding how to produce animals that grow quickly without compromising their welfare or the quality of the final product and with least impact on the environment. One of the primary ways to conduct genetic improvement historically is by using selective breeding. Prior to the advent of modern genetics genetic improvement was conducted using animal pedigree records and phenotypes. Simply put, for meat production larger animals were selected as parents and bred together more frequently to produce larger animals that grew more quickly in subsequent generations. However, the advent of modern statistical genetics allows a greater degree of precision in identifying animals with superior genetic effects; breeding these animals in successive generations leads to an entire breed or line with superior genetic ability to grow quickly in various environments (Kirkpatrick & Heckman, 1989; Hermesch et al., 2003). It must be noted that maintaining animal performance in various environments is one of the main challenges facing animal breeding due to the myriad of interactions between genetics and environment that produces the phenotypes we observe.

## 1.2 GROWTH MODELS IN UK SHEEP PRODUCTION

The methods available model weight or another growth trait as a function of age, often with a set of fixed or random environmental effects. Environmental effects are the sources of relationship between animals that do not come from shared genetic relationship but from similar habitats, rearing environments, or fixed traits such as sex. Fixed environmental effects can be described by discrete characteristics of an animal's habitat such as climatic information, sex, or management group and they remain constant across observations of the trait or interest. Random environmental effects cannot be characterized by fixed variables and are variable across observations meaning animals within a group may respond differently to the same environment. Current methods to map growth and select superior animals for breeding relies on two prominent assumptions: that the mean and covariance structure of the growth trait remains constant with time or age and that each measurement of the growth trait is genetically different from but correlated to all other measurements of the growth trait (Amer et al., 1997; McLaren et al., 2014). Multivariate models are then created to produce estimated breeding values (EBV) at fixed age points which are then used typically to create a selection index (Jones et al., 2004).

This methodology is problematic. It has three main issues: (1) impracticality; (2) neglecting wide random variability in environmental effects; and (3) overparametization. The first issue arises because the current methodology relies on weights taken at specific points in time or within certain windows which may or may not coincide with the normal rearing procedures of the producer (Bullock et al., 1993). Consequently, the model represents a very specific subset of the animal's entire growth cycle rather than the entire growth trajectory (Schnyder et al.,

2001; Lewis and Brotherstone, 2002). On the second issue, environmental fixed effects are fitted to the age or point in time of the weight measurement. This means that the only environmental effect included is the one available at the time the weight was recorded rather than capturing the way environment may impact an animal's growth the entire time it is growing (Hermesch et al., 2003; Jones et al., 2004). Finally, the models run the risk of being overparameterized wherein there are more parameters than can be estimated from the data (Ashfield et al., 2014). This occurs when there is insufficient information in the data to estimate the number of parameters being fitted. This can result in genetic parameters which are not positive definite and ill behaved. In the case of the growth modelling this can cast doubt on the accuracy of the curve the model produces.

### **1.3 GENETIC MODELLING AND SELECTION OF UK SHEEP**

Because this project applies random regression models and genomic models to study growth curves in UK sheep and thereby estimate genotype x environment interactions, it is useful to review the current state of sheep modelling in the UK agriculture industry. Sheep production is an important part of agricultural production and wider economic growth (Amer et al., 2007; Ashfield et al., 2014; Simm et al., 2021). Current research focuses largely on maximizing growth to target market weight while minimizing the amount of time and economic inputs necessary to achieve that. Growth trajectories represent an individual animal as a function that relates the age of that individual to some measure of its size (Kirkpatrick and Heckman, 1989).

Growth models use longitudinal data where the same trait is measured multiple times over an animal's lifetime, where the trait of interest is changing but in a gradual manner over a certain period (Meyer, 2001). Thus, research can estimate the animal's ability to grow quickly. Of course, measuring growth is merely the first step to optimizing growth. Here it should be noted that optimizing growth takes into consideration multiple facets of animal husbandry and market

considerations. There are increased financial and environmental costs associated with feeding and caring for large breeding animals later in life. Thus research into sheep growth seeks to understand how to incorporate fat cover, muscularity at various points in the animal's body, the changes in body condition score across the animal's life cycle and the impact of environmental variance to optimize the economic return farmers receive as well as the overall yield and quality of the sheep carcass at the point of slaughter (Lambe et al., 2004, 2005, 2007; Amer et al., 2007; Connington et al., 1995; Simm et al., 2021). While there may be a benefit for breeding dams to have moderate body condition scores and moderate carcass fat, animals that are too large may be less functionally fit (Lambe et al., 2005). Caring for sick animals creates additional farm cost (Simm et al., 2021). However, it is usually beneficial for animals to grow quickly early in life so long as it does not compromise their health.

Multiple studies have investigated various ways to affect or select for an animal's growth (Lewis and Brotherstone 2002; Hermesch et al., 2003; Jones et al., 2004). Feed efficiency is often a focus because food is one of the major costs in any production system. Therefore, if one could maximize feed intake and nutrient acquisition and minimize time, the cost of food in relation to the final product would decrease (Hermesch et al., 2003). Bio-economic models are also widely used to capture the impact of a production system on an animal's growth. This aims to effectively describe complex systems by accounting for genetic, nutritional, management, and economic factors (Jones et al., 2004). These models provide a good way to estimate the economic impact of genetic changes in traits and to investigate the way changes in non-genetic factors impact the values of these genetic changes (Jones et al., 2004). The use of growth models such as these has improved the efficiency of livestock production. Hermesch et al (2003) have reported a significant change in the efficiency of pig production from identifying and selecting for animals that grow efficiently, which may also select for animals

with lower maintenance requirements. Models that identify and map growth over time seek to account for variation between animals. Jones et al. (2004) conducted a study where the model accounted for variation in rates of growth and fattening, seasonal fluctuations in market prices, feed requirements, and seasonal variation in feed costs.

Growth models have informed entire national production systems in countries across Europe. For example, Irish beef cattle systems are based on maximizing early live weight gain to slaughter using grass-based diets (Ashfield et al., 2014). There is a clear broad focus in production systems in the United Kingdom on growth traits. Historically, this focus on growth traits was realized by breeding the animals in a flock or herd that grew quickly (Simm et al., 2021). This phenotypic selection was imperfect as it neglected wide ranges of environmental variability. However, with the advent of modern genetic science, evidence arose that mature weight could be changed genetically by selection (Bullock et al., 1993). Thus, production systems in the United Kingdom began focusing on genetic selection for growth traits rather than using phenotypic selection or seeking to isolate certain beneficial environmental factors like climatic characteristics or nutrient ration. With increased emphasis on genetic selection, consideration for the ways in which environmental factors affect the trait of interest as well as other traits is important (Bullock et al., 1993). Additionally, genetic selection must be sensitive to environmental variance. Especially in the case of UK sheep production, environments vary widely in terms of resources and management schemes. This may lead to genotype x environment interactions which could negatively impact the efficiency of selection and the performance of flocks (Wilaston, 1987; Maniatis and Pollott, 2002a).

Thus, growth models need to identify genotypes best suited to various environments or genotypes that perform well across different environments. This would allow producers to use

genotype x environment interactions to their benefit (McLaren et al., 2014). There are many environmental and physiological components which growth models seek to account for. For example, in dairy cattle, diet and disease can affect growth through metabolic hormone concentrations but these are very difficult to monitor (Brickell et al., 2009). For sheep, lameness, gastrointestinal worms, or farm system impacts growth (Marshall et al., 1991; Maniatis and Pollott, 2002b; Roeber et al., 2013). Again, statistical models seek to find ways to account for these minute changes in management practice. Jones et al. (2004) published a model which accounted for variation between lambs in rates of growth and fattening, seasonal fluctuations in market prices, feed requirements and seasonal variation in feed costs. Indeed, growth models are used to account for an incredibly wide array of effects that impact growth. Constructing growth models that can produce representations of genetic merit in the form of estimated breeding values (EBV), account for the impact of environmental factors, and address the possible impact on other traits are now of great importance to the economic viability of and animal quality in the UK's agricultural sector.

There is long term economic benefit to using biological growth models that are sensitive to environmental effects. They can be extended to describe conformation and fatness in the carcass over time which can be related to preexisting economic frameworks to establish a profit maximizing slaughter point (Amer et al., 1997). Additionally, knowing the size an animal is projected to be at a certain age can be associated with a wide range of economically important traits (Amer et al., 1997). However, this is a difficult task in practice. Carcass quality characteristics, market prices, and production costs change over time making estimating an animal's optimum slaughter age difficult. Genotype specific slaughter point criteria has been shown to be important in economic comparisons in countries such as Canada (Amer et al., 1997). Despite the proven economic importance of genetic selection, price incentives to

producers in the UK are likely to be based, in the short term, on the operational decisions of producers and finishers (Amer et al., 1997). Additionally, EBV estimates that come from genetic models for UK sheep have limited uptake across the industry which may negatively impact their perceived utility (AHDB, 2021). Amer et al., (2007) noted in their review of the benefits of genetic progress in the sheep and beef industries that the most uncertainty in their models was in sheep projections because the impact of cooperative breeding programs on the commercial sector of the industry was unknown. Further work is also being done to increase the amount of data available for mature animals which will help to improve the predictions' accuracy and hopefully lead to more widescale use of EBV in the sheep industry (AHDB, 2021). Therefore, accurately mapping the ways genetic selection, management factors, and genotype x environment interaction is incredibly important to one of the UK's most economically significant sectors.

#### **1.4 FIXED POINT GROWTH MODELLING**

Longitudinal growth data are frequently analyzed for academic publication by fitting a repeatability model which assumes all records were repeated measurements of a single trait with constant variances (Ptak and Schaeffer, 1993; Guo and Swalve, 1995). Alternative approaches have been to subdivide ranges of ages and consider that range as an individual representation of the trait. Each of the age segments were considered an individual trait and then each of these individual representations of the trait were analyzed in a multivariate model (Meyer, 1999; Legarra et al., 2004). Both methodologies suffer from certain limitations that have already been discussed, such as practicality, ignoring variability, and over-parameterization. Lewis and Brotherstone (2004) identified another issue in their study of sheep growth. The repeated measurements of growth are used as a kind of path from birth to

maturity. Functions are used to describe what happens along the way with relatively few parameters. However, the estimates of these parameters, such as weight, are highly correlated. From a biological point of view this makes sense. If one measures a trait over time in the same individual the measurements will share environmental and genetic sources of variance. Failing to account for this correlation means the model's prediction of genetic value for each animal is ultimately less accurate (Lewis et al., 2004). McLaren (2014) identified another source of concern for the current methodologies. Because repeatability models can only account for the environmental variance at play when the fixed weight measurement is taken, the animals are selected based on the environmental conditions present for a few fixed points in time. When an animal is then used in another location or the same location but under different environmental conditions the producer may obtain unexpected and inconsistent results (McLaren et al., 2014). McLaren et al (2014) found that when models include variables to cluster farms or flocks such as climatic conditions or herd size the resulting EBVs or heritability values could be quite different from those derived from models without these variables.

## **1.5 RANDOM REGRESSION MODEL DEFINITIONS AND APPLICATIONS**

### ***1.5.1 Definition of the Random Regression Model***

The random regression model seeks to remedy these issues by providing a quantitative genetic model for infinite dimensional characters that is sensitive to wide environmental variability, many weight measurements throughout an animal's life, and genetic correlation between records. This methodology in an animal breeding context is a recent innovation pioneered by researchers working on improving the prediction of genetic value in dairy cattle (Lewis et al.,

2004; Schaeffer, 2004; Coffey et al., 2006). The basic structure of the random regression model remains generally constant across various applications (Schaeffer, 2004). It is:

$$y_{ijkn:t} = F_i + g(t)_j + r(a, x, m1)_k + r(pe, x, m2)_k + e_{ijkn:t},$$

Where  $y$  is the  $n$ th observation of the  $k$ th animal at time  $t$  belonging to the  $i$ th fixed factors and the  $j$ th group.  $F_i$  is a fixed effect. These are independent of the time scale being used such as age or months. An example would be a management group effect or flock effect. The next variable,  $g(t)_j$  is a function that accounts for the phenotypic trajectory of the average observation across all animals belonging to the  $j$ th group. The next term  $r(a,x,m1)_k$  is the notation adopted for the random regression function. This version denotes the additive genetic effect for the  $k$ th animal,  $x$  is the vector of time covariates, and  $m1$  is the order of the regression function. The next variable is another regression coefficient for the permanent environmental effect for the  $k$ th animal. Lastly  $e$  is the random residual effect (Schaeffer, 2004).

The random regression model produces coefficients that give rise to covariance functions which describe the covariance structures for the complete range of ages in the data, even ages that are not included in the dataset (Meyer, 1999). Thus, the random regression model can account for a wider variability in growth rate than a repeatability model. The notable difference arises because the repeatability model maps growth as a trend across a subset of individual datapoints while the random regression model represents growth as a function rather than a finite set of measurements (Kirkpatrick and Heckman, 1989; Mrode, 2005). There are three distinctive advantages to the infinite dimensional method. The first is it describes the trait at all points in time during the animal's life rather than a finite number of landmarks. Second, it eliminates the errors in predicting the evolutionary response to selection made by conventional methods that neglect the impact of selection on other traits or other components of the same

trait. Third, it is more efficient at estimating parameters of interest (Kirkpatrick and Heckman, 1989).

This methodology is also better able to capture environmental variability. When the environmental variable changes in a continuous or fluctuating way, such as temperature or grass availability, the genotype's reaction norm is a function consisting of an infinite number of points (Kirkpatrick and Heckman, 1989). The reaction norm depicts the range of phenotypes that can be produced by the genotype given different environments. The infinite dimensional method describes this wide-ranging variability by applying functions to these characteristics rather than fixed points. The first advantage of the infinite dimensional method, interpolating a curve for the whole of an animal's life, is minor considering that some forms of curve fitting can be used to interpolate between the points of a finite dimensional analysis (Kirkpatrick and Heckman, 1989). The second advantage presents itself when predicting how the average value of a trait will change in response to selection (Kirkpatrick and Heckman, 1989). The third advantage, statistical efficiency, presents because the function is given more and more points along the continuum of the character in question which it interprets as a set of results from one model rather than many models working with subsets of the data (Kirkpatrick and Heckman, 1989). The random regression model fits polynomials to the data which have been shown to converge much more quickly to an asymptotic value, which is presumed to be the actual eigenvalue.

This method may account for the spacing of the sample points when the orthogonal polynomials are fitted to the covariance functions whereas the finite dimensional methods exemplified in the repeatability model ignores the spacing and order of the finite points. Here it is important to note that the finite dimensional methodology suffers from two key limitations.

The first is that it cannot accommodate non-linear functions as a representation of the data. It is difficult to validate the assumption that the data would naturally follow a linear relationship. The second issue is the phenomenon of ‘wiggly’ functions. These arise where the function is unable to accurately describe the data due to fewer data points at the extremes of the distribution. The application of random regression modelling is currently expanding. Now most of the research on random regression concerns lactation records in dairy cattle. There are currently relatively few estimates of genetic merit for sheep and other production livestock species (Albuquerque and Meyer, 2001; Meyer, 2002; Nobre et al., 2003; Lambe et al., 2004; Lambe et al., 2005). The random regression methodology must adapt for application to growth because unlike lactation records, animal growth is cumulative over time (Schaeffer, 2004) The key concerns when applying random regression models to growth traits are the number of times an animal’s weight needs to be recorded, when those recordings need to be taken, and what the upper limit on age should be. Additionally, current management practices may introduce accidental bias. Records are kept on only a subset of animals which tend to be managed more directly and those animals which are kept passed standard market weight for breeding tend to be the highest performing animals. These sources of bias must be accounted for to ensure the accuracy of the models. The cost of collecting weights is also practically important for the use of random regression. Now, random regression offers more freedom in this regard for timing of weighing than the repeatability model, but the number of weights is still an important consideration. The more data there is available the more precise the model can be.

### ***1.5.2 Application of Random Regression Models***

Random regression models have also been applied to understand genotype x environment interactions. If such interactions exist, this implies that animals would rank differently in

different environments in terms of performance. The random regression can be applied if the environment can be ordered by for example average herd production or by temperature or by another variable (Schaeffer, 2004). Although research is looking into using random regression models to account for environmental variance, further research is needed to consider higher order polynomials and include fixed curves to account for phenotypic relationships (Schaeffer, 2004).

Random Regression modelling has been successfully used to better capture the variability of an animal's continuous growth. Meyer used a repeatability model and a cubic regression to model weight as a function of age in cattle (Meyer, 1999). The study concluded that the repeatability model did not fit the data used in the study even when the model allowed for heterogenous measurement error variances (Meyer, 1999). This result suggested that the weight of adult cows throughout their lives could not be regarded as repeated measurements of one trait with a constant variance and heritability (Meyer, 1999). In another study live weight of lambs was modelled including additive genetic and permanent environmental effects fitted to a fifth order polynomial and dam effects fitted to a third order polynomial. The models result in variable heritability values across the growth curve. They ranged from 0.09 at 15 days of age to 0.33 at 155 days of age. The model's results suggested that a lamb's live weight early in life is a different trait from its live weight later in life. The genetic variation between one lamb and another was accounted for by fitting random regressions for each animal (Lewis and Brotherstone, 2002). Growth in dairy cows and corresponding genetic merit has also been mapped using random regression models (Coffey et al., 2006).

This model from Coffey et al., (2006) for growth in dairy cows resulted in heritability estimates of 0.53, 0.45, and 0.75 for birth weight, weaning weight, and calving weight respectively. The

confidence intervals for these genetic correlations between the weight traits suggested that the traits were not under the same genetic control. They were 0.79 between birth weight and weaning weight, 0.50 between birth weight and calving weight, and 0.59 between weaning weight and calving weight. This result is evidence that growth is likely to be modelled most effectively by a model that can account for variable genetic variances across a growth curve because the genetic control varies from one weight to the next. Lambe et al. (2006) also found random regression models applied to growth in Scottish Blackface sheep and Texel sheep resulted in different heritability across the growth curve and between breeds. The resulting heritability values ranged from 0.14 to 0.74 in Texel and 0.07 to 0.34 in Scottish Blackface, which suggests that modelling growth requires a model that allows for changing variance structures.

Similarly, Sallam et al (2019) found that heritability, and therefore genetic variability, changed across the growth curve in sheep data. Their study used 16,496 records that included five weights for each individual. The Legendre Polynomial random regression model included third order for a fixed effect of lamb age and permanent environmental effect and a fourth order Legendre polynomial for the animal genetic effect. The heritability estimates were 0.06 for birth weight, 0.42 for weight at nine months, and 0.27 for weight at one year of age. The study compared the Random Regression Model with a multi trait model and found generally lower heritability estimates in the multi trait model. The Random regression model also showed higher accuracy of the predicted breeding values which indicates the random regression model is more suited to analyzing longitudinal growth data in sheep (Sallam et al., 2019). Molina et al (2007), however, found heritability increased with age. In their study with 88,727 individual Spanish Merino lambs, a random regression model with regression of order 3 for the direct and maternal genetic effects and for the individual permanent environmental effect resulted in

heritability values for the direct genetic effect ranging from 0.123 to 0.186. Although the reported heritabilities are different from one study to the next, all these studies reported that heritability changes across the growth curve, which suggests that the random regression model is best suited to model growth in a population of sheep or cattle.

Another study modelled growth in dairy cows by using fixed regression coefficients, the random regression coefficients associated with the animal plus its permanent environment, and the residual error associated with days since calving (Coffey et al., 2003). Coffey et al. (2003) used Legendre polynomials in the random regression. Similarly, Guo et al. (2002) sought to estimate covariance components across multiple lactations and test for heterogeneity of genetic and environmental variance across records. The data were analyzed using random regression models that allowed the covariance between lactation measurements to change gradually with days in milk (Guo et al., 2002). In the case of lactations, the benefit of the random regression model is that it accounts for the shape of an animal's lactation curve through regressions and models short term environmental effects that increase the specificity of the genetic evaluation (Jamrozik et al., 2002). Random regression models have also been used to model energy balance profiles over multiple lactations in dairy cattle. Coffey et al. (2003) found that lactations 1 through 3 could be modelled with moderate accuracy using sinusoidal functions. Beef cattle growth has also been modelled with random regression models. Genetic parameters for a random regression model of growth in Gelbvieh beef cattle was constructed using results from previous studies of random regression models to model growth in this species of cattle (Legarra et al., 2004). This example of random regression found that cubic Legendre polynomials seemed to provide a reasonable fit for the data (Legarra et al., 2004). Clearly random regression models have been used effectively to model growth in sheep and cattle. They offer the unique flexibility required to account for changing heritability values and

variable variance structures across an animal's lifetime. However, current research continues to investigate the performance of these models with few data points, how many weight records are ideal for successful modelling, and the feasibility of using these models with production data rather than data collected for experimental purposes in research environments.

## **1.6 THE AIMS OF THIS THESIS AND CONTRIBUTIONS TO THE PRODUCER**

This project aims to use a random regression model to describe growth more accurately in sheep allowing a wider range of growth recordings to be included thereby producing better selection criteria to allow breeders to choose the best breeding stock. This project will model the genetic and environmental covariance structure of weight in sheep by using random regression models. These models will include genetic effects, maternal genetic effects, and a host of conformation and meat quality data, available from the Ramcompare project of AHDB, to fully capture the way an animal's genetic effect impacts growth and overall animal performance. By incorporating multiple sources of variance and multiple traits, this project may be able to produce EBV that represent animals that do not only grow quickly but also have superior conformation and meat quality traits, in short breeding an overall better animal.

This PhD offers a unique contribution to the current body of research as it applies random regression models to a wide range of growth data on UK sheep breeds and it seeks to produce EBV not only for growth but also for economically important carcass and meat quality traits. The objectives of this research are:

- I. Examine the state of currently available growth data. Determine the data structure and adequacy of data for each sheep breed for random regression modelling. Explore various growth curves and estimate genetic parameters based on the best model.
- II. Investigate the genetic relationships between body weights and between body weight and other non-longitudinal traits.
- III. Investigate the genetic relationships between body weight traits and carcass data. Use growth curve parameters to formulate prediction equations for various carcass traits to identify the best weight and age combinations for maximization of economic value of the carcass trait.
- IV. Examine genotype x environment interactions using a reaction norm model in the context of a random regression model.

Specifically, this research includes data from Suffolk, Charollais, and Scottish Blackface breeds as well as F<sub>1</sub> crosses of the same breeds. It also includes reaction norm models that identify genotype x environment interactions thereby accounting for the specific variability presented by each environment. Use of similar random regression and reaction norm models within national genetic evaluations to produce more accurate EBV for growth could ultimately lead to genetic selection indexes that optimize economic returns for the producer. Ultimately, the models in this PhD will identify genotype x environment interactions, correlations with

carcass traits, and optimized selection criteria for sheep in the UK's agricultural sector. Currently, polynomials used in random regression models can have difficulty modelling commercial data due to data structure. Thus, this study will seek to respond by investigating the effect of different numbers of weight records and variable distributions of records across the age distribution. It will also contribute to evidence on the order and kind of polynomial that best fits growth data. The research undertaken in this PhD will not only offer UK agricultural producers specialized EBV that account for specific environmental variability and desirable production traits but will also contribute to establishing random regression modelling as the superior way to model animal growth.

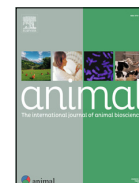
## **CHAPTER TWO: MODELLING GROWTH IN SUFFOLK AND CHAROLLAIS SHEEP POPULATIONS USING RANDOM REGRESSION MODELS AND VALIDATION OF CONSTRAINED POLYNOMIAL CORRELATION VALUES.**

### **2.1 INTRODUCTION**

In this chapter, the research focused on identifying genetic parameters contributing to growth in Suffolk and Charollais sheep. The analysis was initially attempted using random regression models to model growth and estimate subsequent genetic parameters for these two sheep breeds. However, the research revealed that random regression models struggle to model growth and accurately partition variance into genetic, environmental, and residual variances when there are few records into maturity or when data are not normally distributed. The research described here explored an alternative methodology so that random regression models can be used when the data are imperfect or would normally not be appropriate for these models. The methodology was used in both breeds of sheep and genetic parameters were estimated. A validation method was then undertaken to ask if the resulting genetic estimates of potential for growth mirrored the animals' actual growth curves. This chapter offers not only further information on when random regression modelling is appropriate but also investigates solutions to expand the utility of this modelling technique in commercial environments where data are likely to be imperfect.

**2.2 Text Published at:**

**E. McGowan, M. Coffey, G. Simm, R. Mrode,  
Modelling growth in Suffolk and Charollais sheep populations using random regression  
models and validation of constrained polynomial correlation values,  
*animal*, Volume 17, Issue 5, 2023, 100792, ISSN 1751-7311,  
<https://doi.org/10.1016/j.animal.2023.100792>.**



## Modelling growth in Suffolk and Charollais sheep populations using random regression models and validation of constrained polynomial correlation values



E. McGowan<sup>a,b,\*</sup>, M. Coffey<sup>a</sup>, G. Simm<sup>b</sup>, R. Mrode<sup>a</sup>

<sup>a</sup>Scotland's Rural College, Roslin Institute Building, Bush Estate, Midlothian EH25 9RG, United Kingdom

<sup>b</sup>Global Academy of Agriculture & Food Systems, University of Edinburgh, The Royal (Dick) School of Veterinary Studies and The Roslin Institute, Easter Bush Campus, Midlothian EH25 9RG, United Kingdom

### ARTICLE INFO

#### Article history:

Received 31 October 2022

Revised 15 March 2023

Accepted 21 March 2023

Available online 30 March 2023

#### Keywords:

Constrained parameters

Genetics

Heritability

Livestock production

Longitudinal traits

### ABSTRACT

Random regression modelling has been used across multiple animal species to model longitudinal data. The random regression model for growth accounts for the genetic correlation between measures of the same trait over time and the wide environmental variability in growth, but this requires adequate weight records across the age range. However, contemporary management practices in sheep in the United Kingdom generally focus on growing lambs and neglect mature weight recordings. This study examined modelling strategies for growth data in Suffolk and Charollais sheep, provided by the Agriculture and Horticulture Development Board, with polynomial random regression modelling with many early life weight recordings but limited weight recordings in mature animals. Two methods were employed to model the data. In Method A, missing mature weight records were predicted for those animals that did not have a recorded mature weight. The animals were sorted into groups based on the identity of their sires and the year in which the animal was born. Mature weight values were predicted within each group with a multiple regression model. The dataset, including predicted values, was analysed with random regression models using polynomials and simple linear regression for animal and permanent environmental (PE) effects. In Method B, the dataset with missing mature weight records was analysed using a random linear regression animal model with random animal and PE effects. Due to problems of convergence because the parameters were close to the boundary space, fixing the correlation between the intercept and slope of the Legendre polynomial at different levels was investigated. The heritability values resulting from the model with a fixed correlation between intercept and slope parameters at 0.5 for the Suffolk dataset resulted in heritability values ranging from 0.2 to 0.5 from 1 to 619 days of age. Corresponding estimates for the Charollais dataset ranged from 0.18 to 0.49 from 1 to 640 days of age. For the Suffolk data, the genetic correlations ranged from 1.00 to 0.08 between weight at day 1 to weight at day 619, while for the Charollais, the correlations ranged from 1.00 to 0.05 from 1 to 640 days of age. Validation procedures were undertaken using a multitrait approach to examine the estimated breeding values when the correlation between the intercept and slope are fixed at different levels. The results indicated that fixing the correlation at 0.5 gave the most appropriate estimates for the Suffolk and Charollais datasets.

© 2023 The Authors. Published by Elsevier B.V. on behalf of The Animal Consortium. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

### Implications

This research investigates modelling strategies for sheep data with a sparsity of mature weight records relative to early life records. Modelling such data is an important step for geneticists and animal breeders as most sheep production systems do not record mature weights thereby impacting the ability to accurately

model growth over the life course and select for more appropriate growth profiles. The methods used in this research may allow genetic research on a wider range of datasets, therefore improving the efficiency of the sector and contributing to long-term profitability and sustainability.

### Introduction

There are around 30 million sheep and lambs in the United Kingdom (UK) as of 2020 with a production value of £1.3 billion. Employment in the sheep sector contributes about £290 million

\* Corresponding author at: Scotland's Rural College, Roslin Institute Building, Bush Estate, Midlothian EH25 9RG, United Kingdom.

E-mail address: [emutch@sruc.ac.uk](mailto:emutch@sruc.ac.uk) (E. McGowan).

<https://doi.org/10.1016/j.animal.2023.100792>

1751-7311/© 2023 The Authors. Published by Elsevier B.V. on behalf of The Animal Consortium. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

to the UK economy. Sheep production contributes substantially to the UK food supply. The UK is the world's fifth largest producer and third largest exporter of sheep meat. Annually, the UK produces 300 000 tonnes and exports 80 000 tonnes valued at £373.8 million (Agriculture and Horticulture Development Board, 2021). Clearly, the sheep sector is one of the most substantial areas of agriculture in the UK. However, the industry is facing multiple challenges. The global population is increasing and is estimated at 8 billion people (World Population Clock, 2022). Additionally, meat consumption across the world has increased at an even higher rate over the past 50 years and is expected to continue to do so (Vranken *et al.*, 2014). Improving sustainability of production is vital.

Animals that obtain their optimal slaughter weight earlier also emit less greenhouse gas thereby limiting their impact on the environment. Animal science has an important role to play in identifying and selecting animals that grow optimally. Genetic selection has contributed to sheep production by modelling growth and identifying those animals with superior genetic potential to grow optimally and pass on this potential to their offspring. Until recently, repeatability models were used to model growth relying on two prominent assumptions: that the mean and covariance structure of the growth trait remains constant with time or age, and that each measurement of the growth trait is genetically different from but correlated to all other measurements of the growth trait (Amer *et al.*, 1997; McLaren *et al.*, 2014). Multivariate models are then created to produce estimated breeding values (EBVs) at fixed age points which are then used to create a selection index (Jones *et al.*, 2004; Mrode, 2005).

However, this methodology is problematic. It has three main issues: (1) impracticality; (2) neglecting wide random variability in environmental effects; and (3) overparameterisation. The first issue arises because the current methodology relies on weights taken at specific points in time or within certain windows which may or may not coincide with the normal rearing procedures of the producer (Bullock *et al.*, 1993). Consequently, the model represents a very specific subset of the animal's entire growth cycle rather than the entire growth trajectory (Schnyder *et al.*, 2001; Lewis and Brotherstone, 2002). On the second issue, environmental fixed effects are fit to the age or point in time of the weight measurement. This means that the only environmental effect included is the one available at the time the weight was recorded. This makes it impossible to capture the way environment may impact an animal's growth the entire time it is growing (Hermesch *et al.*, 2003; Jones *et al.*, 2004). Finally, the models run the risk of being overparameterized wherein there are more parameters than can be estimated from the data (Ashfield *et al.*, 2014).

There are three distinctive advantages to the infinite dimensional method. The first is it describes the trait at all points in time during the animal's life rather than a finite number of landmarks. Second, it eliminates the errors in predicting the evolutionary response to selection made by conventional methods that neglect the impact of selection on other traits or other components of the same trait. Third, it is more efficient at estimating parameters of interest (Kirkpatrick and Heckman, 1989; Schaeffer, 2004).

The random regression model has been used to model growth in sheep and estimate genetic parameters. Lambe *et al.* (2006) also found random regression models applied to growth in Scottish Blackface and Texel sheep resulted in different heritabilities across the growth curve and between breeds. The resulting heritability values ranged from 0.14 to 0.74 in Texel and 0.07 to 0.34 in Scottish Blackface, which suggests that modelling growth requires a model that allows for changing variance structures. Similarly, Sallam *et al.* (2019) found that heritability, and therefore genetic variability, changed across the growth curve in sheep data. Their study used 16 496 records that included five weights for each individual. The Legendre Polynomial random regression model

included third order for a fixed effect of lamb age and permanent environmental effect and a fourth order for the animal genetic effect. The heritability estimates were 0.06 for birth weight, 0.42 for weight at nine months, and 0.27 for weight at one year of age. The study compared the random regression model with a multitrait model and found generally lower heritability estimates in the multitrait model. The random regression model also showed higher accuracy of the predicted breeding values which indicates the random regression model is more suited to analysing longitudinal growth data in sheep (Sallam *et al.*, 2019). The growing body of research suggests heritability values from 0.10 to 0.40 for growth in most breeds of sheep (Fossco and Notter, 1995; Jannoune *et al.*, 2015; Lalit *et al.*, 2016; Zeljic *et al.*, 2019).

The key concerns when applying random regression models to growth traits are the number of times an animal's weight needs to be recorded, when those recordings need to be taken, and what the upper limit on age should be. Additionally, current management practices may introduce accidental bias. Records are kept on only a subset of animals which tend to be managed more closely and those animals which are kept past standard market weight for breeding tend to be the highest performing animals. These sources of bias must be accounted for to ensure the accuracy of the models. The cost of collecting weights is also practically important for the use of random regression. Now, random regression offers more freedom in the timing of recording than the repeatability model, but the number of weights is still an important consideration (Meyer, 1999). The more data that are available, the more precise the model can be.

Additionally, most random regression models are used successfully in highly managed flocks with equal or high numbers of weight records across the age distribution. Despite the fact that performance of the random regression model is not as well tested in environments with unequal spacing or sparse weight records, there are a few studies that have successfully tested random regression models in actual production environments (Sarti *et al.*, 2015; Arthy *et al.*, 2020). Most production systems in the UK do not regularly record weight records throughout a sheep's life and there is particular sparsity towards maturity. As indicated earlier, the use of a random regression model offers producers more flexibility on when to weigh their animals and on the availability of EBVs at various time periods to suit their production system (Coffey *et al.*, 2006). The preliminary results from this paper have been published in a poster presented at the British Society of Animal Science conference 2022 (Mutch *et al.*, 2022). Thus, this study aimed to investigate the use of random regression models with data collected from production environments across the UK which have comparatively few records in maturity relative to the number of individuals in the dataset. Although mature weight is generally not a focus for genetic research, it is difficult to accurately model the growth curve, and hopefully identify the genetic and environmental factors contributing to optimized early growth rates or slaughter weights, without a mature weight acting as an anchor point for the model. This research may expand the utility of the random regression model to datasets which previously would not be feasible.

## Material and methods

### Description of data

This study was conducted with two datasets provided by the Agriculture and Horticulture Development Board (AHDB); one with Suffolk sheep and their F1 crosses, and the other with Charollais sheep and their F1 crosses. The Suffolk dataset had records from 14 870 Suffolk sheep. It included multiple weight records

for each animal, resulting in a total of 42 244 records. Four possible weight records were identified for the analysis including birth weight, weaning weight (about 9 weeks), scan weight (about 21 weeks), and mature weight which was defined as any weight after 180 days. There were 14 870 animals with a birth weight, 14 823 with a weaning weight, 11 548 with a scan weight, and 94 with a mature weight. From birth weight to mature weight, the number of animals with records began at 100%, dropped to 99.6% for weaning weight, 77.7% for scan weight, and 0.63% for mature weight. The Charollais dataset contained 148 334 individuals with 353 958 records with weights available at birth, weaning, scanning and maturity. There were 148 334 animals with a birth weight, 148 321 with a weaning weight, 146 331 with a scan weight, and 921 with a mature weight, and these details are summarised in Table 1. Initial attempts to use the whole data for genetic parameter estimation resulted in problems with convergence of model. In an attempt to achieve a balanced distribution of weight at the various ages and hence enhance convergence, only individuals with all four weight recordings were kept for use in the model for genetic parameter estimation. This resulted in a total of 3 685 records from 921 animals being used.

The data were collected from farms across the UK over a duration of 22 years from 1997 to 2019 from multiple production environments including lowland, upland, and hill flocks. Summaries of the fixed effects of interest for both datasets are included in Table 2. The Suffolk dataset included both male and female animals born in 33 seasons, a contemporary grouping for animals born within the same season of the same year and raised in 30 flocks. Animals were classified into four management groups. The Charollais dataset included only females from 16 different flocks which were born in 22 seasons and were classified into seven different management groups.

There was minimal intentional selection used during the data collection phase as the data were collected from production environments operating as they would normally operate. Both rams and ewes are included in the datasets. However, all animals with a mature weight record are ewes due to the normal production procedure wherein most male sheep are usually slaughtered upon reaching the appropriate slaughter weight for the breed while ewes are often kept into maturity for breeding purposes. Thus, these data are representative of general Suffolk and Charollais populations in the United Kingdom.

### Statistical analysis of data

Data were subjected to quality measures by first excluding records with missing birth weight using the “filter” function in R. Additionally, the mean and variance structure of each weight variable were analysed and then records which were more than three standard deviations away from the mean were excluded from the dataset using the “filter” function in R. For the Suffolk dataset, animals were excluded from the analysis dataset if they had fewer than two weight records. For the Charollais dataset, there were sufficient animals with all four weight records; thus, the Charollais analysis dataset only included animals with all four weight records. Each of the variables was investigated for content and then used in a linear model for each weight record using the “glm” function in R (Rstudio Team, 2012). The linear models were used to identify those variables which were significant as fixed effects for each weight record. A version of the data was then organised for random regression analysis. In the case of Charollais sheep, all animals had a weight record at the four different ages but for the Suffolk, the number of weights per animal ranged from one to four recorded at various ages. This was done using the “filter”, “select” and “nrow” functions in R (Rstudio Team, 2012).

### Mature weight prediction

This study used two methodologies to model the data. The first method (Method A) can be considered a prediction method where the missing mature weight records were predicted and then analysed. The second method (Method B) modelled the data with only the recorded mature weight records; no records were predicted. Method A used the same Suffolk dataset described previously to predict missing mature weights. The missing mature weights were predicted for the dataset based on those animals that did have a recorded mature weight. To make these predictions, the data were filtered in R using the “filter” function to include only animals with a mature weight.

The sires and birth years of these animals were identified and a prediction model for each of the sire groups and each of the year groups was made. In the first instance, predictions were made within sire groups because it was thought that the mature weights were likely influenced by the impact of sire’s genetic merit on the weight, growth rate, or carcass characteristics of offspring (Held

**Table 1**  
Summary of records for each weight recording point in the Suffolk and Charollais datasets.

Dataset	Birth Weight Count	Weaning Weight (avg 9 weeks) Count	Scan Weight (avg 21 weeks) Count	Mature Weight (above 180 days of age) Count
Suffolk	14 870 (100%)	14 823 (99.7%)	11 548 (77.7%)	94 (0.63%)
Charollais	148 334 (100%)	148 321 (99%)	146 331 (98.6%)	921 (0.006%)

Abbreviations: avg = average.

**Table 2**  
Summary of variables for Suffolk and Charollais datasets.

Dataset	Sex No	Season No	Flock No	Management Groups No	Mean BW (kg)	Mean WW (kg)	Mean WW Age (days)	Mean SW (kg)	Mean SW Age (days)	Mean MW (kg)	Mean MW Age (days)
Suffolk	F: 6 392	33	30	4	M:5.3 ± 1.2	M:30.5 ± 8.6	M:65.4 ± 11.7	M:57.5 ± 13.4	M:145.8 ± 16.4	M:67.9 ± 7.2	M:218.5 ± 8.7
	M: 8 478				F:5.0 ± 1.2	F:27.8 ± 7.8	F:65.2 ± 11.7	F:48.2 ± 11.0	F:144.1 ± 16.6	F:75.9 ± 12.0	F:479.7 ± 124.5
Charollais	F: 3 685	22	16	7	5.2 ± 0.9	26.3 ± 5.4	62.9 ± 10.1	46.35 ± 7.6	137.4 ± 18.5	75.94 ± 12.4	522.4 ± 91.1
	M: 0										

Abbreviations: F = female, M = male, BW = birth weight, WW = weaning weight, SW = scan weight, MW = mature weight.

et al., 1997; Elizalde et al., 2019). However, due to the small number of animals with a mature weight record, it was impossible to predict the mature weight within sire groups for every animal. Therefore, prediction models were made within year groups as animals born in the same year are likely to be of similar genetic merit. The prediction model was based on a multiple regression model with the following equation:

Mat weight = birthweight + weaning weight + scan weight.

The multiple regression model assumes there must be a linear relationship between the outcome variable and the independent variables. In this case, this would mean there is a linear relationship between the mature weight and the three other weights. It also assumes that residuals are normally distributed. The model also assumes that the independent variable, the three weights, are not highly correlated with each other to create any major collinearity problems in predicting mature weight. This assumption is generally tested by using variance inflation factor values. The average inflation factor value for birth weight was 1.91, 2.65 for weaning weight, and 2.24 for scan weight. The model variance will inflate by these values compared to a model without collinearity.

The prediction modelling was done using the “lm” function in base R (Rstudio Team, 2012). These models were then used to predict the mature weights for the entire dataset using the “predict” function in RStudio within sire groups and year groups for animals without recorded mature weights (Rstudio Team, 2012). This methodology was performed only with the Suffolk dataset due to the genetic model from Method A failing to converge and the fact that the Charollais dataset did not include any individuals with missing mature weight records.

#### Genetic modelling

The availability of mature weights in 0.63% of the data implies that the implementation of a multitrait model is not feasible for this data set. Therefore, a random regression model offers the only opportunity to model growth along the whole age trajectory in the dataset for these two breeds. Initial investigations of random regression models for these both data sets included a genetic maternal component in the models. However, these did not converge, and this may be attributed to the data structure or data size that makes it difficult to accurately partition this component. However, the inclusion of maternal effect is being investigated in further studies due to its potential importance for early growth. Genetic modelling began after a three-generation pedigree was formed in R using a “match” function and a full pedigree provided by AHDB (Rstudio Team, 2012). In Method A, the Suffolk dataset with predicted mature weights were analysed in a random regression model (Model 1) with Legendre polynomials of order 1 for animal and PE effects. The fixed effects for this model were sex, season of birth, flock, and management group.

The structure of the model was as follows:

$$Y_{ijkn.t} = F_i + g(t)_j + r(a, x, m1)_k + r(pe, x, m2)_k + e_{ijkn.t} \quad (1)$$

where  $y$  is the  $n$ th weight observation of the  $k$ th animal at time  $t$  belonging to the  $i$ th fixed factors and the  $j$ th group;  $F_i$  represents fixed effects consisting of sex of animal (two levels), season of birth (33 levels), flocks (30 levels), and management groups (four levels). The next variable,  $g(t)_j$ , is a function that accounts for the phenotypic trajectory of the average observation across all animals belonging to the  $j$ th group. The next term  $r(a, x, m1)_k$  is the notation adopted for the random regression function, denoting in this case the additive genetic effect for the  $k$ th animal,  $x$  is the vector of

age in days, and  $m1$  is the order of the regression function which is 1 in this model;  $r(pe, x, m2)_k$  is PE effect for the  $k$ th animal and lastly,  $e$  is the random residual effect.

Model 1 failed to converge. It was determined this was due to there being too few mature weight records to accurately predict mature weight records for the entire dataset. At this point, Model 1 was no longer considered to be a candidate model for the data because it was determined that it was unrealistic to predict mature weights for the dataset from such a small number of recorded mature weights. Method B used an animal model without any prediction of missing records. For Suffolk, a random regression model (Model 2) with Legendre polynomial of order 1 for the random animal effect, random PE effects and fixed effects of sex season of birth, a fixed regression on age, flock, management group, and PE effect was fit.

This model had the following structure:

$$Y_{ijkn.t} = F_i + g(t)_j + r(a, x, m1)_k + r(pe, x, m2)_k + e_{ijkn.t} \quad (2)$$

where terms in Eq. (2) were defined as in Eq. (1).

Although other studies on random regression modelling for growth in sheep report a Legendre polynomial of order 3–5 as the best fitting, in this study, a Legendre polynomial of order 1 was considered to be a sufficient model because the models failed to converge when the order of the Legendre polynomial was increased above 1 or set to 0 (Arthy et al., 2020; Jannoune et al., 2015; Lewis and Brotherstone, 2002).

However, for the model fitted, there were inflated genetic parameters which were not biologically probable, because the correlation between the intercept and slope parameters of the Legendre polynomial was close to unity. When these parameters are approaching unity, this indicates that the model is unable to estimate a value for each of the weight recording points. In the case of the Suffolk model, this is likely due to the lack of records in maturity which could cause the model to be unable to partition the variance between each weight recording point. The biological expectation is that the weight records should be correlated but not identical; thus, the correlation between the parameters of the Legendre polynomial was fixed at three levels to indicate minimal correlation, medium correlation, and high correlation. This allowed for the comparison of the resulting genetic parameters from these three levels of correlation. The resulting genetic parameters should be more accurate when a model with a more biologically plausible correlation value is used compared to the correlation close to unity because the latter deviates both from the biological expectation and from the assumptions of the model. Therefore, the correlation between the intercept and slope was fixed at three different levels of 0.2, 0.5 and 0.7. These three levels of correlations investigate the impact of fixing this correlation at a low, medium and high values. Genetic parameters were estimated from the estimates of the random regression coefficients. All genetic analyses were conducted in ASREML (Gilmour et al., 2016).

For the Charollais dataset, only Method B was used initially in which a model (Model 3) was fit with Legendre polynomial of order 1 for the random animal effect, random permanent environmental effect and the following fixed effects: fixed regression on age, season of birth, and flock was fit.

This model had the following mathematical structure:

$$Y_{ijkn.t} = F_i + g(t)_j + r(a, x, m1)_k + r(pe, x, m2)_k + e_{ijkn.t} \quad (3)$$

where terms were as defined in Eq. (1) but with  $F_i$  consisting of the fixed effects season of birth (22 levels), fixed linear regression on age and flocks (16 levels).

Again, a Legendre polynomial of order 1 was deemed to be the most preferable because the models failed to converge above an order of 1. Model 3 resulted in biologically improbable heritability

values initially, again due to the correlation between the intercept and slope parameters of the Legendre polynomial approaching unity. Again, the three fixed correlations between the intercept and slope coefficients were investigated and genetic parameters were estimated from the results. All genetic analyses were conducted in ASREML (Gilmour et al., 2016).

#### Correlation validation

Attempts were made to validate the genetic parameters obtained from the random regression models with the correlation between the intercept and slope parameters fixed to 0.2, 0.5 and 0.7. The predictive ability of the genetic parameters from the three fixed levels was examined by using a multitrait prediction of EBVs for birth weight, weaning weight, scan weight and mature weight. This was done by implementing the multi trait model using the genetic and residual variances and covariances obtained from the models when the correlation between intercept and slope was fixed at three different levels, fitting the fixed effects as listed in Model 2 (Suffolk) and Model 3 (Charollais), with animal and PE effects fitted as random effects. Additionally, a univariate fixed effect model was implemented for birth weight, weaning weight, scan weight and mature weight individually. These models included only fixed effects and no random animal effect. This analysis was aimed at obtaining records for each of these traits corrected for all fixed effects apart from the random animal effect such that the only components in these corrected records were the genetic merit of the animals and residual error.

The predictive ability and accuracy of the breeding values from the multitrait analysis, using the genetic parameters when the correlations were fixed at three levels, were estimated by correlating the EBVs with corrected records, regression of the corrected records on EBVs and estimating bias for each weight trait for the 94 animals that had observations for all four weight traits. These calculations were all performed in Excel. It should be noted here that bias is represented as mean and standard deviations of the difference between corrected records and EBV.

The validation method applied in this study is similar to the validation approach employed in genomic studies where genomic breeding values are compared with corrected yield in terms of correlations, regression and computations of bias and mean square error (Pryce et al., 2012; Kalaldehy et al., 2021). Also, the method used to compute bias has been used in other studies (Schenkel et al., 2002; Sood et al., 2020). The validation method used in this study examined the ability of the genetic parameters and breeding values generated from the random regression model when the correlation between the intercept and the slope is fixed to predict the performance (corrected for all fixed effects) of animals with actual observations in the validation set. The approach may suffer for the limitations that the validation animals were part of the dataset

used in estimating the random regression parameters. However, such a limitation will equally affect the three scenarios considered and may therefore have no major influence in terms of choosing the most appropriate fixed correlation. In addition, given the data structure, these animals could not be excluded from the model as these were the animals with data along the complete trajectory of the age in the study.

Attempts were made to examine the impact of fixing the correlation between the intercept and the slope using dairy cattle data in which a random regression model has been successfully implemented (Mrode et al., 2021). The model in Mrode et al (2021) was modified such that Legendre polynomial of order 1 was fitted for both animal and PE effects to mimic the model for this sheep study. The correlation between the intercept and slope was 0.1 after convergence of the model. The analysis was then repeated with the correlations fixed at 0.2, 0.5 and 0.7 and estimates of genetic parameters compared with the original method with a correlation of 0.1. Estimates of heritabilities were then generated for 10 stages of lactation ranging from day 4 to day 364. The results from these analyses are in the discussion section for this paper as the main focus of this study is the sheep data.

## Results

#### Genetic parameters for the Suffolk and Charollais breeds

Method A for the Suffolk dataset failed to converge, meaning any resulting genetic parameters are unlikely to be accurate. Therefore, results have not been presented. The results from Method B, Model 2, for Suffolk are shown in Table 3. These results are presented for the model using the fixed correlation of 0.5 between intercept and slope parameters. Prior to fixing the correlation, Model 2 resulted an additive genetic variance of 0.02 for BW, 13.58 for WW, 68.68 for SW, and 990.92 for MW. This corresponded to heritabilities of 0.02, 0.49, 0.69, and 0.76, respectively. The genetic correlations were BW-WW = 0.46, BW-SW = 0.02, BW-MW = 0.01, WW-SW = 1.0, WW-MW = 1.0, MW-SW = 1.0. Table 3 shows that the heritability values resulting from the fixed correlation of 0.5 were between 0.20 and 0.50 for ages ranging from 1 to 619 days of age. The table also includes genetic (0.1–1.00) and phenotypic (0.03–0.60) correlations which follow the anticipated trend of decreasing as the difference between ages increases. The PE effect accounted for 15.41 kg<sup>2</sup> out of a total variance of 2 984.41 in the 0.2 model, 0.52 kg<sup>2</sup> out of 54.04 in the 0.5 model, and 1.26 kg<sup>2</sup> out of 61.45 in the 0.7 model.

The results from Model 3, the Charollais model, are shown in Table 4. These are also reasonably consistent with the published estimates between 0.10 and 0.40. The heritability values range from 0.18 to 0.49. The table also includes genetic (0.05–1.00) and phenotypic (0.01–0.44) correlations which follow the anticipated

**Table 3**

Genetic (below diagonal) and phenotypic (above diagonal) correlations among several ages in days (along the top and left side) with heritabilities on the diagonal from the random regression model for Method B<sup>1</sup> for Suffolk.

Day	1	30	60	120	180	200	244	460	530	619
1	0.23	0.22	0.21	0.19	0.17	0.17	0.15	0.08	0.05	0.03
30	1.00	0.22	0.21	0.20	0.18	0.17	0.16	0.10	0.08	0.05
60	0.98	1.00	0.21	0.20	0.18	0.18	0.17	0.12	0.10	0.08
120	0.92	0.95	0.98	0.20	0.19	0.19	0.19	0.16	0.15	0.14
180	0.82	0.87	0.91	0.98	0.20	0.20	0.20	0.21	0.21	0.20
200	0.78	0.83	0.88	0.96	1.00	0.20	0.21	0.22	0.22	0.22
244	0.68	0.74	0.80	0.91	0.98	0.99	0.22	0.25	0.26	0.26
460	0.26	0.34	0.43	0.61	0.77	0.81	0.88	0.37	0.60	0.42
530	0.17	0.26	0.35	0.54	0.70	0.75	0.84	1.00	0.43	0.46
619	0.08	0.17	0.26	0.46	0.64	0.69	0.78	0.98	1.00	0.50

<sup>1</sup> Method B refers to the Methodology in which only recorded mature weights were used in Model 2 and then, parameters were constrained to a correlation of 0.5.

**Table 4**

Genetic (below diagonal) and phenotypic (above diagonal) correlations among several ages in days (along the top and left side) with heritabilities on the diagonal from the random regression model for Method B<sup>1</sup> for Charollais.

Day	1	31	60	120	182	200	255	460	530	640
1	0.22	0.21	0.20	0.18	0.16	0.16	0.14	0.07	0.05	0.01
31	1.00	0.20	0.20	0.18	0.17	0.16	0.15	0.09	0.07	0.04
60	0.98	1.00	0.19	0.18	0.17	0.17	0.16	0.11	0.09	0.07
120	0.92	0.95	0.98	0.18	0.18	0.18	0.17	0.15	0.14	0.13
182	0.81	0.86	0.91	0.98	0.19	0.19	0.19	0.19	0.19	0.19
200	0.78	0.83	0.88	0.96	1.00	0.19	0.20	0.21	0.21	0.20
255	0.65	0.72	0.78	0.89	0.97	0.98	0.21	0.24	0.25	0.25
460	0.25	0.34	0.42	0.60	0.76	0.80	0.90	0.35	0.37	0.40
530	0.16	0.24	0.34	0.53	0.70	0.74	0.85	1.00	0.40	0.44
640	0.05	0.14	0.23	0.43	0.62	0.67	0.79	0.98	0.99	0.49

<sup>1</sup> Method B refers to the Methodology in which only recorded mature weights were used in Model 3 and then, parameters were constrained to a correlation of 0.5.

trend of decreasing as the difference between ages increases. Thus, the results for the Suffolk and Charollais models are similar and reasonably similar to published estimates for heritability of weight.

#### Validation results for the Suffolk breed

The results for the Suffolk validation procedure are shown in Table 5. For the Suffolk dataset, the correlations from the fixed correlation of 0.2 and 0.5 indicate similar correlations between corrected record and EBV for the four traits while the correlations were generally higher at 0.70. The higher correlations between EBVs from the parameters obtained from a correlation of 0.7 with corrected record may be due to the high heritability estimates (0.26–0.63) obtained at this fixed correlation compared to the heritability estimates of (0.23–0.50) at the fixed correlation of 0.5. The regression coefficients for 0.7 are closest to 1, suggesting a regression of unity wherein the EBV and corrected records values are the most similar, for birth weight, scan weight, and mature weight. The regression coefficients are generally similar for 0.2 and 0.5 across all four weights points. The bias between 0.2 and 0.5 is generally similar for each weight trait except for birth weight where the bias for 0.20 is much further away from 0 when it is compared to the bias for 0.5. A value of 0.7 produces biases that are closer to 0, except the bias for mature weight comparing to a value of 0.5. The validation results seem to indicate that 0.7 is the correlation of choice for the Suffolk dataset. However, the estimates of genetic correlation among weight traits were negative in some cases. Thus, in addition to the validation results, comparison of genetic parameters to published estimates was employed as an additional measure to select the most appropriate estimates. This leads to the selection of a fixed correlation of 0.5 as the most correlation given the validation results and the genetic parameters resulting from this fixed correlation.

#### Validation results for the Charollais breed

The results for the validation for the Charollais dataset are given in Table 6. The bias was generally better, that is, closest to 0, for the fixed correlation of 0.5 across all four traits. While the correlations were higher for the fixed correlation at 0.2 for scan weight, the correlations tended to be similar across the fixed correlations for the other three weights traits. No clear pattern can be observed for the regression coefficients as they showed under or over predictions for the various traits at each of the three fixed correlation levels. However, given the lowest bias for the fixed correlations at 0.5, and the similar or slightly higher correlations, compared to the fixed correlation of 0.2 or 0.7, the results from the fixed correlations at 0.5 were chosen as most appropriate for the Charollais dataset.

#### Discussion

The random regression model has been useful in modelling growth in several animal populations. However, research continues to investigate ways to overcome the constraints of the model when it is applied to field data from more realistic production environments. The results from this study contribute to a growing body of research that establishes the need for many records across an animal's entire growth cycle to accurately model growth using random regression modelling. In this study, with limited mature weights in less than 0.63% of the Suffolk data and 0.006% of the Charollais data initially, the implementation of a multitrait model is not feasible for this data set. Therefore, a random regression model offers the only opportunity to model growth along the whole age trajectory in the dataset for these two breeds. However, this study also used a multi trait validation on constrained estimates of genetic parameters from random regression models as a potential strategy to model growth in animals with few or

**Table 5**

Validation results for fixed correlations between intercept and slope at three different levels for the Suffolk breed.

Item	BW			WW			SW			MW		
	0.20	0.50	0.70	0.20	0.50	0.70	0.20	0.50	0.70	0.20	0.50	0.70
Cor <sup>1</sup>	0.28	0.27	0.36	0.74	0.71	0.81	0.47	0.45	0.55	0.70	0.72	0.89
Reg <sup>2</sup>	0.04	0.05	0.50	0.76	0.68	1.31	0.33	0.32	0.50	0.74	0.73	0.88
Bias <sup>3</sup>	-2.46	-1.99	-0.80	-2.22	-2.45	-1.14	-3.69	-3.70	-3.03	-4.16	-4.29	-4.47
SD of bias	2.46	2.02	1.61	3.45	3.59	3.50	4.95	5.02	4.69	6.60	6.32	5.20

Abbreviations: BW = birth weight, WW = weaning weight, SW = scan weight, MW = mature weight, CY = corrected record, EBV = estimated breeding value.

<sup>1</sup> Cor is the correlation between CY and EBV.

<sup>2</sup> Reg is the regression coefficient.

<sup>3</sup> Bias is calculated as the mean difference between the CY and EBV.

**Table 6**

Validation results for fixed correlations between intercept and slope at three different levels for the Charollais breed.

Item	BW			WW			SW			MW		
	0.20	0.50	0.70	0.20	0.50	0.70	0.20	0.50	0.70	0.20	0.50	0.70
Cor <sup>1</sup>	0.05	0.36	0.32	0.70	0.68	0.68	0.87	0.78	0.74	0.88	0.87	0.87
Reg <sup>2</sup>	0.01	0.18	0.18	0.53	1.36	1.33	0.63	2.14	1.74	1.32	1.86	1.25
Bias <sup>3</sup>	-6.46	-0.19	-0.35	-0.46	-0.08	-0.41	-3.27	-0.35	-0.90	-1.46	-0.04	-1.66
SD of bias	61.52	1.37	1.27	3.35	2.71	2.75	3.35	4.58	4.68	5.23	6.06	5.24

Abbreviations: BW = birth weight, WW = weaning weight, SW = scan weight, MW = mature weight, CY = corrected record, EBV = estimated breeding value.

<sup>1</sup> Cor is the correlation between CY and EBV.<sup>2</sup> Reg is the regression coefficient.<sup>3</sup> Bias is calculated as the mean difference between the CY and EBV.

unequally distributed weight recordings throughout the growth cycle. These results will hopefully allow more data to be used in random regression modelling which will expand the potential benefits of genetic selection and growth modelling to a larger portion of the UK sheep population.

#### Genetic parameters for the Suffolk breed

Method A is not useful because the model failed to converge in the first place suggesting that although the predicted parameters are biologically plausible, the model was not able to accurately predict a growth curve for the dataset. The results from constraining the estimates of the genetic parameters in Method B for the Suffolk are similar to current literature estimates. However, prior to fixing the correlation, the results from Method B have a highly inflated heritability ranging up to 0.76 at the top of the age range at 619 days. This heritability most likely reflects the very small dataset available to predict mature weight with only 0.63% of the animals having a mature weight record. The prevalence of edge effects in which the random regression model struggles to accurately fit a growth curve at the extreme ends of the age distribution or where there are limited records has been documented in other studies (Lambe et al., 2006). However, from a biological standpoint, one would not expect the genetic heritability of the growth trait to increase so sharply from 0.02 to 0.76 as it does in Method B prior to fixing the correlation.

One of the potential reasons the initial genetic parameters are inflated in Method B may be due to the inability of the model to partition the genetic variation adequately along the trajectory. The inability of the model to partition the genetic variance could be due to data size and structure in this study, given that random regression models have been successfully fitted in some sheep studies. The biological expectation is that the measurements are correlated to but not identical to each other. However, fixing the correlation may introduce some bias into the model's predictions. Thus, the validation was an attempt to verify if the model predicted accurate genetic values for each animal.

When comparing the accuracy of the constrained parameter levels for Suffolk, the better predictive ability produced from a value of 0.7 as indicated by some of the regressions because the regression coefficients for 0.7 are closer to 1 for some of the weights. The corrected record from the fixed effect models represents the animal's genetic value for the growth trait and residual error. They therefore provide an independent variable to assess the accuracy of predictions when the correlation is fixed at different levels. If the constrained parameter models are accurately predicting the genetic merit of the animals for the growth trait, then these EBVs should have a high correlation as the actual genetic merit obtained from the fixed effect models and should have high predictive ability. Thus, when the corrected records are regressed on the EBVs for each fixed correlation model, the most accurate model will have a regression closest to unity represented by a

value of 1. However, it must be noted that the Suffolk dataset had very few (94) animals. The validation results do not clearly distinguish which correlation is the most accurate as the results were similar for each constrained model. Thus, the validation results can be used to support a constrained estimate of genetic parameters only as long as the genetic parameters from the constrained model are biologically reasonable.

The parameters obtained from a correlation of 0.5 were considered more biologically appropriate compared to the results from the fixed correlation of 0.7 for the Suffolk breed. This was due to two main considerations. First, there were a small number of animals available for the validation, at only 94 animals with all four weight records, which may have caused similarity between the correlations and regression at 0.5 and 0.7. Second, the range of heritability values from a correlation of 0.5 were more consistent with the published estimates, which range from 0.10 to 0.40. Lewis et al. (2002) reported a heritability value of 0.37 for mature size in sheep, 0.38 for growth rate, and 0.31 for weight at 150 days of age. Other studies report values between 0.10 and 0.40 (Fossceco and Notter, 1995; Jannoune et al., 2015; Lalit et al., 2016; Zeljic et al., 2019).

#### Genetic parameters for the Charollais breed

The heritability values displayed in Table 4 are in keeping with the findings of other studies. Additionally, the Charollais heritability values are very slightly lower than those for Suffolk, remaining slightly more consistent with current literature estimates. However, this difference is minimal and likely due to a larger number of records in maturity present in the Charollais dataset rather than any difference in model accuracy. For the Charollais dataset, the high correlations cannot be due to a sparsity of mature weight records because all the animals included in the dataset have mature weight records. However, the animals that possessed mature weight records were all female animals from similar production environments. Additionally, it is possible that because all weight records were taken at different ages for each animal, the model was unable to converge because it had to predict weights for a wide range of individual ages. Further studies could investigate condensing the age points into weeks or months rather than days to aid convergence.

The results for the validation for the Charollais dataset more strongly support the selection of the 0.5 fixed correlation. The bias was, generally, closest to 0 for the fixed correlation of 0.5 across all four weight traits. Although the regression coefficients did not indicate a regression of unity consistently for 0.5 across all four weight traits, there is no clear pattern for either of the other fixed correlations. Given the lowest bias at a fixed correlation of 0.5 and the generally higher correlations, 0.5 was chosen as the best fixed correlation. Additionally, the heritability values for 0.5 are more consistent with published estimates, making 0.5 the more appropriate choice for Charollais sheep. As previously discussed, the cur-

rent literature estimates for growth in sheep range from 0.1 to 0.4 (Fossco and Notter, 1995; Jannoune et al., 2015; Lalit et al., 2016; Zeljic et al., 2019).

#### The validation procedure

The validation procedure was used to determine if the constrained estimates of genetic parameters predicted accurate genetic trends for each animal. The validation helps to determine estimates that are usable and that the estimates from the chosen levels, in this case the genetic parameters from the constrained correlation level of 0.5, corresponds to the animal's expected genetic merit for growth represented by the corrected yields. In this procedure, the corrected yields should be highly correlated to the EBVs from the most accurate model. This is because, ideally the model is producing EBVs that are identical to the corrected records. Additionally, as stated, the regression coefficient for the corrected record regressed on the EBV should indicate unity for the most accurate correlation level. The bias estimates should be close to 0 as they are calculated as the mean and standard deviation of the difference. The difference for the correct correlation level should be minimal.

This approach is rather pragmatic, but it yields estimates that are practically useful given the limitations of these datasets. The fixed levels and their validation will be further studied in a beef data set that contains complete records. The limitations of the sheep dataset will be mimicked in the analysis of beef data by omitting records in the beef dataset. Additionally, as more data become available for these particular sheep breeds, the parameters can be further validated even though the estimates are practically useful for now.

The impact of fixing the correlation between the intercept and slope in the dairy models showed that the mean difference between heritabilities at the 10 stages of lactation as a deviation from the original model were  $-0.009$ ,  $-0.04$  and  $-0.08$  for correlations fixed at 0.2, 0.5 and 0.7, respectively. However, mean difference was very close to that for the fixed correlation of 0.5, if estimates from the original model were deviated from estimates averaged across the three fixed correlations. Differences in estimates of genetic correlation followed the same pattern. The results showed that as the difference between the actual and the fixed correlations increases, there was a corresponding increase in estimates of genetic parameters. However, the differences were, on average, not very large. In addition, the fixed correlation of 0.5 seems to give estimates that are similar to estimates averaged across the fixed regression. In the case of the sheep data, the high correlation between the intercept and the slope resulted in estimates that were not biologically feasible. It seems that fixing the correlation at 0.5 provided average genetic parameters that reasonably span the age trajectory for the data set.

#### Conclusions

The availability of mature weights in only 0.63% of the Suffolk dataset and 0.006% of the Charollais dataset implies that the implementation of a multitrait model is not feasible using these data sets. Therefore, a random regression model offers the only opportunity to model the growth data for these two breeds. The approach of constraining the genetic parameters from the parameter space and validating these genetic parameters offers a pragmatic method to obtain genetic parameters for BW across the growth trajectory that are consistent with literature estimates. However, as more data become available, the method may need to be further investigated. Further studies using beef data are currently underway to investigate the impact of limited weight

records at various ages of the growth trajectory on genetic parameters when fitting a random regression using Legendre polynomials.

#### Ethics approval

Not applicable.

#### Data and model availability statement

The data/models were not deposited in an official repository. The data/models that support the study findings are available from the authors upon request.

#### Author ORCIDS

**Emma McGowan:** <https://orcid.org/0000-0002-1155-6501>.

**Raphael Mrode:** <https://orcid.org/0000-0003-1964-5653>.

**Mike Coffey:** <https://orcid.org/0000-0003-4890-6218>.

**Geoff Simm:** <https://orcid.org/0000-0002-4801-903X>.

#### Author contributions

**Emma McGowan:** Conceptualisation, Methodology, Writing – Original Draft. **Raphael Mrode:** Validation, Resources, Writing – Review and Editing, Supervision. **Mike Coffey:** Supervision, Writing – Review and Editing. **Geoff Simm:** Supervision, Writing – Review and Editing.

#### Declaration of interest

None.

#### Acknowledgements

The authors acknowledge the Agriculture and Horticulture Development Board (AHDB) for access to their data collected on Suffolk and Charollais sheep provided for this research which is part of a PhD at SRUC and the University of Edinburgh funded by AHDB.

#### Financial support statement

This research is part of a PhD project funded by the support of the Agriculture and Horticulture Development Board, grant number: 054-2018 AHDB Postgraduate Studentship

#### References

- Agriculture and Horticulture Development Board, 2021. UK sheep facts and figures. Retrieved on 5 September 2022, from <https://ahdb.org.uk/uk-sheep-facts-and-figures>.
- Amer, P.R., Emmans, G.C., Simm, G., 1997. Economic values for carcass traits in UK commercial beef cattle. *Livestock Production Science* 51, 267–281.
- Arthy, V., Venkataramanan, R., Sivaselvam, S.N., Sreekumar, C., Balasubramanyam, D., 2020. Suitability of random regression models for growth of Madras Red sheep under a field performance recording system. *Small Ruminant Research* 193, 106260.
- Ashfield, A., Wallace, M., Prendiville, R., Crosson, P., 2014. Bioeconomic modelling of male Holstein-Friesian dairy calf-to-beef production systems on Irish farms. *Irish Journal of Agricultural and Food Research* 53, 133–147.
- Bullock, K.D., Bertrand, J.K., Benyshek, L.L., 1993. Genetic and environmental parameters for mature weight and other growth measures in Polled Hereford cattle. *Journal of Animal Science* 71, 1737–1741.
- Coffey, M.P., Hickey, J., Brotherstone, S., 2006. Genetic aspects of growth of Holstein-Friesian dairy cows from birth to maturity. *Journal of Dairy Science* 89, 322–329.

- Elizalde, H.F., Carson, A.F., Munoz, C., 2019. Effects of sire genotype on lamb performance at weaning in extensive sheep systems. *Animal* 13, 213–220.
- Fossco, S.L., Notter, D.R., 1995. Heritabilities and genetic correlations of body weight, testis growth and ewe lamb reproductive traits in crossbred sheep. *Animal Science* 60, 185–195.
- Gilmour, A.R., Gogel, B.J., Cullis, B.R., Thompson, R., 2016. *ASReml User Guide Release 2.0*. VSN International Ltd, Hemel Hempstead, UK.
- Held, J., Slyter, A. L., Read, B., Long, B., 1997. Effect of Sire Selection on Lamb Growth and Carcass Traits. South Dakota Sheep Field Day. Proceedings and Research Reports, 12 June 1997, Brookings, SD, USA, Paper 5.
- Hermesch, S., Kanis, E., Eissen, J.J., 2003. Economic weights for feed intake in the growing pig derived from a growth model and an economic model. *Journal of Animal Science* 81, 895–903.
- Jannoune, A., Boujenane, I., Falaki, M., Derqaoui, L., 2015. Genetic analysis of live weight of Sardi sheep using random regression and multi-trait animal models. *Small Ruminant Research* 130, 1–7.
- Jones, H.E., Amer, P.R., Lewis, R.M., Emmans, G.C., 2004. Economic values for changes in carcass lean and fat weights at a fixed age for terminal sire breeds of sheep in the UK. *Livestock Production Science* 89, 1–17.
- Kalaldehy Al, M., Swaminathan, M., Gaundare, Y., Joshi, S., Aliloo, H., Strucken, E.M., Ducrocq, V., Gibson, J.P., 2021. Genomic evaluation of milk yield in a smallholder crossbred dairy production system in India. *Genetics Selection Evolution* 53, 73.
- Kirkpatrick, M., Heckman, N., 1989. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *Journal of Mathematical Biology* 27, 429–450.
- Lalit, Malik, Z.S., Dalal, D.S., Dahiya, S.P., Patil, C.S., Dahiya, R., 2016. Genetic analysis of growth traits in Harnali sheep. *Veterinary World* 9, 128–132.
- Lambe, N.R., Navajas, E.A., Simm, G., Bünger, L., 2006. A genetic investigation of various growth models to describe growth of lambs of two contrasting breeds. *Journal of Animal Science* 84, 2642–2654.
- Lewis, R.M., Brotherstone, S., 2002. A genetic evaluation of growth in sheep using random regression techniques. *Animal Science* 74, 63–70.
- Lewis, R., Emmans, G., Dingwall, W., Simm, G., 2002. A description of the growth of sheep and its genetic analysis. *Animal Science* 74 (1), 51–62.
- McLaren, A., Lambe, N., Morgan-Davies, C., Mrode, R., Brotherstone, S., Conington, J., Morgan Davies, J., Bunger, L., 2014. Characterisation of terminal sire sheep farm systems based on a range of environmental factors: a case study in the context of genotype by environment interactions using Charollais lambs. *Animal* 8, 867–876.
- Meyer, K., 1999. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. *Journal of Animal Breeding and Genetics* 116, 181–205.
- Mrode, R., 2005. *Linear Models for the Prediction of Animal Breeding Values*. CABI Publishing, Wallingford, UK.
- Mrode, R., Ojango, J., Ekine-Dzivenu, C., Aliloo, H., Gibson, J., Okeyo, M.A., 2021. Genomic prediction of crossbred dairy cattle in Tanzania: A route to productivity gains in smallholder dairy systems. *Journal of Dairy Science* 104, 11779–11789.
- Mutch, E., Mrode, R., Coffey, M., Simm, G., 2022. Modelling growth in Suffolk sheep with insufficient mature weight records for polynomial random regression models. Poster presented at 78th Annual British Society of Animal Science Conference, 12–14 April 2022, Nottingham, UK, 18.
- Pryce, J.E., Arias, J., Bowman, P.J., Davis, S.R., MacDonald, K.A., Waghorn, G.C., Wales, W.J., Williams, Y.J., Spelman, R.J., Hayes, B.J., 2012. Accuracy of genomic predictions of residual feed intake and 250-day body weight in growing heifers using 625,000 single nucleotide polymorphism markers. *Journal of Dairy Science* 95, 2108–2119.
- RStudio Team, 2012. *RStudio: Integrated Development for R*. RStudio, PBC, Boston, MA, USA. URL <https://www.rstudio.com/>.
- Sallam, A.M., Ibrahim, A.H., Alsheikh, S., 2019. Estimation of genetic parameters and variance components of pre-weaning growth traits in Barki lambs. *Small Ruminant Research* 173, 94–100.
- Sarti, F.M., Lasagna, E., Giontella, A., Panella, F., Pieramati, C., 2015. The use of a random regression model on the estimation of genetic parameters for weight at performance test in Appenninica sheep breed. *Italian Journal of Animal Science* 14, 310–315.
- Schaeffer, L.R., 2004. Application of random regression models in animal breeding. *Livestock Production Science* 86, 35–45.
- Schenkel, F.S., Schaeffer, L.R., Boettcher, P., 2002. Comparison between estimation of breeding value and fixed effects using Bayesian and empirical BLUP estimation under selection on parents and missing pedigree information. *Genetics Selection Evolution* 34, 41–59.
- Schnyder, U., Hofer, A., Labroue, F., Künzi, N., 2001. Genetic parameters of a random regression model for daily feed intake of performance tested French Landrace and Large White growing pigs. *Genetics Selection Evolution* 33, 635–658.
- Sood, S., Bhardwaj, V., Kaushik, S., Sharma, S., 2020. Prediction based on estimated breeding value using genealogy for tuber yield and late blight resistance in auto-tetraploid potato (*Solanum tuberosum* L.). *Heliyon* 6, e05624.
- Vranken, L., Avermaete, T., Petalios, D., Mathijs, E., 2014. Curbing global meat consumption: Emerging evidence of a second nutrition transition. *Environmental Science & Policy* 39, 95–106.
- World Population Clock, 2022. *Worldometer current world population*. Retrieved on 8 September 2022, from <https://www.worldometers.info/world-population/>.
- Zeljic, K., Stanojevic, D., Bogdanovic, V., Mekic, C., Ruzic-Muslic, D., Maksimovic, N., Stojiljkovic, N., 2019. Estimation of growth traits heritability coefficients in sheep. *Biotechnology in Animal Husbandry* 35, 347–355.

## 2.3 DISCUSSION

In this chapter, random regression modelling was applied to commercial data and methodologies were explored which can be used when initial modelling techniques fail. It also posed further questions that then inspired subsequent work on where the limit of the random regression model lay, in terms of its ability to handle sparse or unevenly distributed growth records. There is evidence in other studies suggesting that data structure can impact genetic evaluations with more records per animal improving genetic estimates (Maniatis and Pollott, 2003). The commercial environment could be responsible for the modelling difficulty. The variability in environmental conditions or the lack of uniform data collection techniques could also contribute to the difficulty. Thus, further research was planned on flocks in research environments with uniform data collection methods and equal weight recordings throughout their life. The research sought to investigate the differences between modelling commercial and research flocks to identify how this can impact genetic evaluations and how modelling can overcome these difficulties. The next chapter will investigate random regression modelling on data from a research flock compared to data from commercial flocks.

Additionally, at the time the paper was published a maternal effect had not been successfully included in the models. This could have also accounted for the modelling difficulty. However, a maternal effect has since been included in the Suffolk models. The maternal effect was included as a random effect and as a random regression with LP of order 1. In the latter case, the correlation had to be constrained to 0.5 again to achieve convergence. In the model with maternal effect fit as a random regression, the maternal effect accounted for 0.17 % of the total variance. For the model with maternal trait as a random effect, the maternal effect accounted for 0.02% of the total variance. The heritability for the growth trait with maternal effect as a random regression model is 0.32; for the model with maternal effect as a random effect the

heritability is 0.32. The heritability for the growth trait in the original model ranged from 0.23-0.50. The addition of the maternal effect has an identical impact on the heritability when included as a random regression and as a random effect which suggests that there is no difference in the applicability of either method in this case. Additionally, the similarity of the heritability with or without a maternal effect suggests that in this case the maternal effect does not alter the genetic variance or the heritability of the growth trait. However, it has been noted in other studies that including maternal additive and environmental effects can significantly improve the fit of genetic models and produce more reliable genetic parameters (Maniatis and Pollott, 2002b; Maniatis and Pollott, 2002c; Conington et al., 1995). It must be noted that the publication mentions using cattle data to mimic the characteristics of the sheep data used in the analysis. The intention was to model growth in cattle and then reduce records to assess the impact on convergence. However, this was not possible due to only 33% of the animals in the cattle dataset having complete parentage records. Additionally, the variances expressed in the Results section of the chapter should all be expressed with the unit of  $\text{kg}^2$ .

Overall, this chapter has sought to identify genetic parameters contributing to growth in Charollais and Suffolk sheep. Additionally, it developed alternative methods to apply random regression modelling to sparse and unevenly distributed datasets. The following chapter will seek to expand the models used in this chapter to include both growth and carcass traits in sheep to calculate indexes that not only capture the animal's growth but also its carcass potential and corresponding profitability. The lessons learned in this chapter about record number and data distribution will be carried into the next chapter to help to achieve model convergence and accurate genetic estimates.

## **CHAPTER THREE: MODELLING GROWTH CURVES IN IRREGULAR SHEEP DATA AND THE RELATIONSHIP BETWEEN GROWTH AND CARCASS TRAITS**

*Much of this chapter published at: Mutch, E., Mrode, R., Coffey, M., Simm, G., 2022., Using Random Regression Models to Model Growth and Environmental Effect in Sheep with Irregular Data, Proceedings of the 12<sup>th</sup> World Congress Applied to Livestock Production July 2022, Rotterdam, Netherlands, 723.*

### **3.1 INTRODUCTION**

Breeding animals for desirable traits is a practice as old as domestication. Animal breeders' top priorities currently includes contributing to a sustainable and efficient food supply for the growing human population whilst using less resource. Thus, animal breeding uses genetic analysis to identify and breed animals that have genetic potential to grow quickly. However, growth is not the only important factor in determining the genetic value of a particular animal. This chapter investigates the use of random regression models to model growth and genotype x environment interactions in commercial flocks and the relationship between carcass and growth traits.

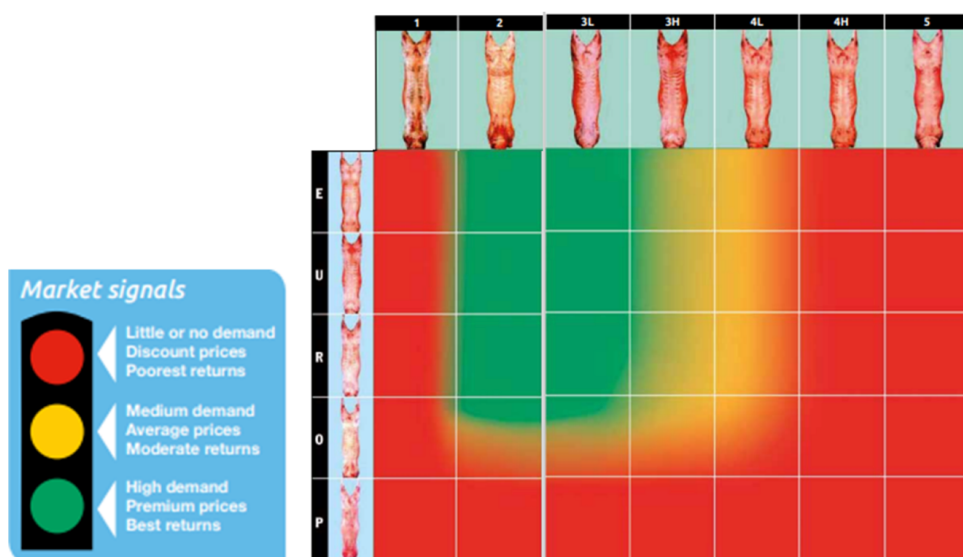
#### ***3.1.1 Random Regression Modelling with Unevenly Distributed Data***

Random regression models represent growth as a function rather than a finite set of measurements, which allows accounting for variability in the rate of growth and in environmental effects throughout an individual's life. However, random regression models can be difficult to fit to real world data which may be unbalanced or sparse. Although traits such as a growth curve can be modelled with a polynomial or linear function, a specific dataset can have variable rates of change or have distinct measurements that are highly correlated among themselves. When measurements are too highly correlated, it can then be difficult to validate the assumption that every measurement of the growth trait is correlated but distinct from the other measurements.

Secondly, a large number of records are required for the model to converge. Random regression models are often used on lactation data in dairy cattle where test day records follow a well-defined pattern compared to growth data; especially in less intensive sheep production systems where data may be sparse. This chapter examines the performance of random regression models in datasets that may have more variability because the data is recorded in commercial environments. Using a commercial dataset and a dataset collected on a research flock allows this study to compare the applicability of the random regression model in these two contexts.

### 3.1.2 Including Carcass Traits in Genetic Models

Particular features of lamb carcasses are financially rewarded at the point of sale to the abattoir. These features are identified in the EUROP grid. The EUROP grid was established in the 1980s by EU legislation to predict meat yield from a combined scale of conformation and fat class scores.



**Figure 3.1:** The EUROP Grid used for sheep carcass grading (Farm Advisory Service, 2024).

There are a possible 56 different carcass classifications in the EUROP grid with the highest premiums paid for carcasses that fall within the green shaded area of the diagram above. Although there is debate on how well the EUROP system meets consumer preferences, the system is used across the UK and determines the premiums and penalties that producers receive for their animals (Bonny et al., 2016). The EUROP grid confirms that speed of growth is not the only trait of interest to sheep breeders and producers. Rather, incorporating conformation and fat class into breeding decisions can have a positive impact on the quality of the animals as well as the profitability of commercial sheep systems. Selection to decrease fat can accelerate muscle growth (Prache, et al., 2022). Breeds like the Texel have been selected for increased muscling. Genetic selection in this breed has resulted in increased muscling, carcass weight, and decreased fat (Prache, et al., 2022). Clearly fat class can be identified and manipulated through selective breeding techniques. Additionally, carcass conformation has a well-documented relationship with growth. The most important factor in determining the conformation of a carcass is age. The carcass becomes thicker and more compact as the animal grows and ages (Prache, et al., 2022). Thus, to optimize the conformation, animals must reach the optimum weight and fat class at maturity before muscles begin to atrophy with advanced age (Prache, et al., 2022).

This study will further investigate the applicability of random regression modelling to sheep data by looking at the relationship between growth and carcass traits.

### **3.2 MATERIALS AND METHODS**

The project was conducted with two datasets. The first was provided by the Agriculture and Horticulture Development Board (AHDB). It consisted of mixed breed animals included in the

Ramcompare (RC) project which sought to identify the genetic merit of key sires currently used in UK sheep breeding. The dataset consisted of approximately 48,614 individuals with a total of 104,874 weight records from the F1 crosses from UK terminal sires crossed on typical UK maternal breeds; for example Texel, mule ewes including Welsh, Scottish and North Country, Lleyn etc. This dataset included four possible weight recordings with an accompanying age for each individual sheep. The animals were weighed at birth, weaning (approximately 9 weeks of age), and at scanning (approximately 21 weeks of age) and a fourth weight which was defined as any weight taken above 180 days of age for both males and females. The weaning weight recorded at nine weeks is young compared to the national average, 12-16 weeks, but this age was the calculated average mean for lamb age at weaning in this particular dataset (AHDB, 2021). The dataset was constrained to only those animals that had at least a birth weight and one other weight recorded. There were 37,692 animals that had a birth weight. 31,069 animals had a weaning weight and a birth weight, 26,633 animals had a scan weight and a birth weight and 1,266 animals had a fourth weight and a birth weight. The RC dataset used in this study was collected across multiple farms in the UK over a duration of four years from 2016 to 2020. The data were collected in various production environments including lowland, upland, and hill systems. Summaries of the fixed effects of interest are included in Table 3.1. The RC dataset included both male and female animals born into 13 different flocks but moving through 23 different flocks, meaning that there were 23 different flocks in which weaning weights, scan weights and mature weights were recorded. The animals were classified into 3 different breed categories.

**Table 3.1:** Display of variables in the RC dataset.

1. The avg is the mean value displayed with standard deviation.

<b>Variable</b>	<b>Description</b>	<b>Levels</b>
<b>Flock id</b>	Flock animals lived in	23
<b>Sheep id</b>	Animal identification number	48,614
<b>Sex</b>	Animal sex	Female: 24,629 Male: 16,958 Castrated: 7,027
<b>Sire id</b>	Animal's sire	443
<b>Dam id</b>	Animal's dam	14,499
<b>Age</b>	Age of the animals in days	1-197 Avg 56.66 ± 44.08
<b>Weight</b>	Weight of the lambs in kg	1 – 68.50 Avg 21.18± 12.43
<b>Muscle depth mm</b>	Ultrasound muscle depth measured in mm	7.50 – 35.70 Avg 11.58 ± 11.39
<b>Fat Depth mm</b>	Ultrasound fat depth measured in mm	0.17 – 11.87 Avg 1.03 ± 1.31

The second dataset included records from a research flock of Scottish Blackface Sheep (SBF) kept by Scotland's Rural College (SRUC) and included 23,256 animals. Each of these animals had four weights recorded. The first was at birth, the second recording was taken at marking, the third weight was recorded at weaning and the fourth was defined as any weight taken after 180 days of age for both males and females. Summaries of the fixed effects of interest are included in Table 3.2. The dataset included both male and female animals born in 20 different years and weight recordings taken in four different seasons per year. The animals were

maintained in the same flock but were divided into 15 different grazing groups. The lambs were raised to grow to a prescribed slaughter weight at which point they were removed from the population. Thus the weights taken after the second weight are not all taken at the same point in the age trajectory because the age at which they are recorded is dictated by growth rate rather than achieving a specific age.

**Table 3.2:** A description of the variables in SBF.

<b>Variable</b>	<b>Description</b>	<b>Levels</b>
<b>Animid</b>	Animal identification number	23,256
<b>Sex</b>	Animal sex	M: 3969 F:19287
<b>Line</b>	Genetic line* of the animal	3
<b>Weight</b>	Weight of the lambs in kg	1.9 – 60.20 Avg 22.76 ± 14.13
<b>Graze code</b>	Grazing group	15
<b>Age</b>	Age of the animals in days	1 – 376 Avg 101.5 ± 91.85
<b>Month</b>	Month in which the weights were recorded	12
<b>YMD</b>	A combination of year month and day at which the weight was recorded	903
<b>Season</b>	Season of weight recording	4
<b>Year Season</b>	Year and season when the weight was recorded	75

- 1 The genetic line of the animal was the purpose for which they were bred; within the flock there were replacement lines meant to be kept for breeding and animals meant to be sent away for slaughter.
2. The Avg is the mean value displayed with the standard deviation.

The various factors defining time including month, YMD, Season and Year Season are included to offer a diversity of possible factors to include in the model to represent the time across which the growth cycle occurs. The random regression model is able to model fixed effects differently for each measurement across the time period given in the model. The different indicators of time allow for comparing different levels of specificity on the time variable which can allow for comparing how well the model fits fixed effects relevant to the exact moment of weight recording compared to a month or season or year of several weight recordings. The fixed effects may have little variability across season but could have variability across a month or exact date. At the outset of modelling this is an unknown but including the different levels of specificity for time allows the modelling to account for the variability that is introduced by the differences in fixed effects in each moment of weight recording across the entire duration of the growth cycle.

There was minimal intentional selection used in the RC dataset because the data were collected from standard commercial environments. Here minimal selection is referring to the fact that the animals were not intensely managed or specifically selected to be in the project based on their growth performance. The data collected from the SBF flock were collected from all animals regardless of their genetic line or involvement in other studies thus removing potential bias from the animal's being intentionally bred for specific trials. Although both rams and ewes are included in both datasets, all animals with a mature weight in the RC dataset are ewes due to the normal production measures of slaughtering males while retaining some females for breeding purposes. Three generation pedigrees were used for genetic analyses in ASREML for both datasets.

Quality control measures were used on both datasets. First, any animals without a birth weight and one other weight record were removed from the RC dataset. This was done using the ‘filter’ function in R. The mean and variance structure of each weight were explored in both datasets. Two versions of the data were then constructed. One version contained all available weight records while the other removed any datum that fell outwith three standard deviations of the mean for each of the individual weight traits. For the RC dataset, animals were excluded if they had fewer than two records while the SBF dataset only included animals with all four weight records available. Each variable was investigated and then used in a linear model for weight using the ‘glm’ function in R (Rstudio Team, 2012). These models allowed significant fixed effects to be identified. The datasets were organized for random regression analysis. For the RC dataset, the number of records per animal could range from two to 4 while the SBF dataset contained only animals with 4 records. This was done using the “filter”, “select” and “nrow” functions in R (Rstudio Team, 2012).

### ***3.2.1 Phenotypic curves.***

Initially, several curves were fitted at the phenotypic level to determine how several functions might best fit the distribution of the data before including the genetic component. The datasets were analyzed using Excel with polynomial regression models of several orders (1 to 6) with weight fitted as a function of age. Spline functions were also fitted in R using the “stats” package and the “rm. Smooth” function in R studio. A Gompertz function was also tested in Excel. These methods were applied to both datasets to initially investigate their distributions. The quantity and size of the full datasets were, at times, impossible to use when they exceeded the computation capacities of Excel or R studio.

### ***3.2.2 Genetic analysis.***

Genetic analysis was conducted on both the RC and SBF datasets using the random regression model. All genetic analyses were conducted in ASRMEL (Gilmour et al., 2016). In both datasets, the aim of genetic modelling was to obtain genetic parameters for the growth trait throughout the distribution of ages and to improve upon the accuracy of EBVs provided by multi-trait genetic modelling. Modelling began for the SBF dataset after three generation pedigrees were formed in R using a “match” function (Rstudio Team, 2012). The random regression model for the SBF data was constructed consisting of the fixed effects: birth year, sex, genetic line, birth rearing rank, dam’s age, grazing code, a fixed regression on age, month of recording, a combination of year month and day of recording, season of recording, and a combination of year and season of weight recording. A random animal effect with Legendre polynomial of order 1 and a permanent environment effect was fitted to the full dataset for SBF using ASREML (Gilmour et al 2016). The low polynomial order fitted in the two studies was only chosen after higher polynomial orders were modelled and convergence could not be achieved. This may be due to the limiting data structure involved in the studies.

Similarly, a random regression model was fitted for the RC data which included fixed effects of breed, flock id, sex, birth type, and a fixed regression on age. The random effects were animal effects modelled by Legendre polynomial of order 1 and a permanent environmental effect using ASREML. In both models fitted for the breeds, initial investigations that included

a genetic maternal component was examined but these models did not converge. This could be due to the variable number of weight records in the RC dataset. However, non-convergence of the maternal effect has been observed throughout this study and the previous study when the correlation between the intercept and slope for the LG was close to the parameter boundary space at 1. This could also be due to the limited data structure in this study.

The structure of the RC model was as follows:

$$Y_{ijkn:t} = F_i + g(t)_j + r(a, x, m1)_k + r(pe, x, m2)_k + e_{ijkn:t} \quad (3.1)$$

Where  $y$  is the  $n$ th weight observation of the  $k$ th animal at time  $t$  belonging to the  $i$ th fixed factors and the  $j$ th group;  $F_i$  represents fixed effects consisting of breed (3 levels), flock id (23), sex of the animal (3 levels) and a fixed regression on age. The variable  $g(t)_j$  is a function that accounts for the phenotypic trajectory of the average phenotypic observation across all the animals belonging to the  $j$ th group. In this case that average phenotypic observation is the weight trait. The next term  $r(a, x, m1)_k$  is the notation used for the random regression function, denoting in this case the additive genetic effect for the  $k$ th animal,  $x$  is the vector of age in days and  $m1$  is the order of the regression function. For this model that order is 1. The next variable,  $r(pe, x, m2)_k$  is the PE effect for the  $k$ th animal and lastly,  $e$  is the random residual effect.

The structure of the SBF model was as follows:

$$Y_{ijkn:t} = F_i + g(t)_j + r(a, x, m1)_k + r(pe, x, m2)_k + e_{ijkn:t} \quad (3.2)$$

The terms in Eq (2) were defined as in Eq (1). However, in this case  $F_i$  represents different fixed effects which are described above. These include sex (2 levels), line (3 levels), birth rank (10), dam's age (1-7), grazing group (15 levels), and a fixed regression on age.

Heritabilities and genetic correlations among several weights at different ages were generated from the analysis. The genetic parameters from the model were inflated and not biologically probable. It was observed that the correlation between the intercept and slope of the Legendre polynomial was close to unity indicating that the model is unable to estimate a value for each of the weight recording points. In the RC dataset, this is likely due to a lack of records into maturity which could prevent the model from appropriately partitioning the variance between each weight recording point (McGowan et al., 2022). The biological expectation is for each weight to be correlated to but not identical to the other weights, and so the correlation was constrained at three levels to reflect a minimal correlation, medium correlation and high correlations. This allowed for a comparison between the resulting genetic parameters from the different correlations applied. The genetic parameters should be closer to reality when a model with a more biologically representative correlation value is used compared to the correlation close to unity because the latter is not the biological expectation nor does it meet the assumptions of the model (McGowan et al., 2022).

Therefore different correlations were fixed at three different levels, were examined and the genetic parameters compared for biological plausibility. The correlations were fixed at 0.2, 0.5, and 0.7 for the SBF dataset in Model 2 and 0.2, 0.5 and 0.9 for the RC dataset in Model 1. The correlations were fixed to represent a low, medium and high correlation. Two different correlations represented the high correlation in these two models because these were the respective initial correlations between the parameters of the Legendre Polynomial when the models were run without the constrained correlation value. Thus, a correlation of 0.7 for the

SBF and 0.9 for the RC most accurately reflected a high correlation in that dataset. Genetic parameters were calculated from each model and compared against published estimates of heritability for growth in sheep around 0.1 -0.4 (Fossceco and Notter, 1995; Jannoune et al., 2015; Lalit et al., 2016; Zeljic et al., 2019). In both cases the model at a correlation of 0.5 produced genetic parameters that were biologically plausible and consistent with literature estimates. The current literature estimates were used as indicators of successful modelling because the study aimed to understand if the random regression model could offer a similar genetic evaluation while also offering greater statistical efficiency and more flexibility for the recording of the phenotype. Thus, values that were wholly different from published literature estimates would indicate that the models provided a different evaluation which was not the aim of the study.

### ***3.2.3 Relationship Between Growth and Carcass Traits for Scottish Blackface Sheep***

#### ***3.2.3.1 Selection Index***

A selection index was initially used as a preliminary investigation into the relationship between the growth trait and carcass traits which included conformation, fat class, carcass weight, slaughter weight and condition score. The selection index was only calculated for the Scottish Blackface because only this dataset included carcass traits. In the case of condition score the multi trait models described below did not converge and therefore it was not possible to construct the selection index for this carcass trait. The conformation, carcass weight and fat class were recorded at the time of slaughter. The slaughter weight was originally assumed to be a separate weight from the fourth

weight because the value recorded for these weights differed. The original dataset used for analysis did not include a slaughter date so the analysis was performed identifying slaughter weight and fourth weight as separate weight recordings. However, upon subsequent investigation into the data collection methods, it became clear that for the majority of males and some females in the dataset the fourth weight and the slaughter weight were recorded within 2 to 3 weeks of each other meaning the weights could not be assumed to be independent weights. However, for this analysis they were included as independent weights due to the information available at the time of analysis. The selection index calculated for slaughter weight therefore shows that the fourth weight and slaughter weight are highly correlated but this is likely due to them being identical rather than a close genetic relationship between two independent recordings of the weight trait.

The index was calculated using the standard approach and using the standard selection index equation (Simm and Dingwall, 1989) as:

$$I = P^{-1} G \text{ where } P^{-1} \text{ is the inverse of the } P \text{ matrix.} \quad (3.3)$$

A G and R matrix consisted of the variances and covariances among the four body weight traits and one carcass trait in turn. The G matrix consisted of the genetic variance, the R matrix the residual variances and the P matrix consisted of the phenotypic variances. These variances and covariances were taken from multi trait animal models that modelled the four weights as individual trait along with one carcass trait in series of bivariate analyses. The carcass trait was fitted as a continuous numeric variable for slaughter and carcass weights. The fat class was included as a factor with

five numeric levels where 1 corresponds to fat class 1, 2 corresponds to fat class 2, 3 corresponds to fat classes 3L and 3H, 4 corresponds to fat classes 4L and 4H, and 5 corresponds to fat class 5. The conformation was included as five numeric levels with 1 corresponding to E, 2 corresponding to U, 3 corresponding to R, 4 corresponding to O, and 5 corresponding to P. A summary of the carcass traits can be found in Table 3.3

**Table 3.3** A summary of Carcass Traits for the SBF Dataset

<b>Variable</b>	<b>Description</b>	<b>Mean</b>	<b>Standard Deviation</b>
<b>Slaughter Weight</b>	Pre-slaughter live weight	40.55 kg	4.480
<b>Carcass Weight</b>	The weight of the slaughtered carcass	17.39 kg	2.430
<b>Fat Class</b>	Numeric indicator of EUROP grid fat class	3.025	0.8180
<b>Conformation</b>	A numeric code indicating the EUROP grid conformation score	3.024	0.5730
<b>Condition Score</b>	A numeric indicator of body condition ranging from 1-5 with the ideal value being 3	3.015	0.1024

The accuracy of the selection index was then calculated using the following equation:

$$R_{IH} = (b'Pb/a'Ca)^{1/2} \quad (3.4)$$

Here  $R_{IH}$  is the correlation between the selection objective and the selection criteria which represents the accuracy of the index; the closer the correlation is to one the higher the accuracy is for the selection index. This is equal to the square root of  $b'$  as the transposed vector of  $b$  values which are the weighing factors or index coefficients by

which the phenotypic measurements are multiplied in the selection index; multiplied by P, the matrix of phenotypic variances and covariances for the weights in the selection index, and by b the vector of b values. This is divided by a'Ca which is equal to the genetic variance of the carcass trait in this case.

Finally, the difference in accuracy between the full selection index involving all 4 body weights and the carcass trait compared to the accuracy when one of the body weights was dropped from the index was computed. This provided an indication of the contribution each body weight made to the overall index. The relative contribution of each body weight to the accuracy of the index was calculated as the accuracy of the reduced index divided by the full selection index.

### ***3.2.3.2 Random Regression Analysis of Carcass Traits in SBF***

In addition to the selection index, bivariate random regression models were used to estimate heritability, genetic correlation and phenotypic correlation for body weights from one day old to 376 days old and one of the following carcass traits: fat class, conformation score, slaughter weight, condition score and carcass weight.

To estimate genetic parameters for both the carcass trait and the weight trait, a bivariate random regression model was constructed for each of the carcass traits. The model structure was as follows:

$$Y^1_{ijkn:t}, Y^2 = F_i + g(t)_j + r(a, x, m1)_k + r(pe, x, m2)_k + e_{ijkn:t} \quad (3.5)$$

The terms in equation (3.5) are as defined in equation (3.1). However, this model is a bivariate random regression model meaning there are two y variables in each model. The carcass traits were included in the same way they were included in the models for the selection index without being adjusted for any covariates. These models included the following fixed effects: sex( 3 levels), genetic line ( 3), season (4), and month (11). Other fixed effects which were included in the dataset were attempted including age, season and year and grazing group but these resulted in the model failing to converge and were therefore removed to ease convergence and obtain genetic values for the models in question.

The correlation between the slope and the intercept of the random regression was constrained for these models due to the findings from the previous growth models for SBF requiring a constrained correlation. Here the order of the random regression Legendre Polynomial was 1.

### **3.3 RESULTS**

#### ***3.3.1 Phenotypic curves***

A summary of the results for the various phenotypic curves fitted is presented in Table 3.3. The polynomial regression had R squared values above 0.80 for both the RC and SBF datasets. The R squared values for the SBF were slightly higher indicating a slightly better fit. Spline functions were selected based on R squared values with the best for RC containing 80 knots and the best for the SBF containing 100 knots. The Gompertz functions for both datasets converged and resulted in equations for both models. The phenotypic comparison was used to

assess if any of the functions would not converge or be unable to model the data which would help to offer a visual assessment of the shape of the growth curve and determine if one shape was more consistent with the data. For example, the Gompertz model uses a sigmoidal curve while the linear regression of order 1 would plot a straight line. All of the models converged which did not offer clear direction on the shape of the growth curve.

### **3.3.2 Genetic analysis.**

Model 1 with a correlation of 0.5 between the intercept and slope was identified as the best model for RC based on the biological plausibility of the parameters. The results produced by this model are presented in Table 3.4. This shows that the heritability values from the fixed correlation of 0.5 ranged between 0.24 to 0.39 for an age range of 1 to 197 days of age. The table presents genetic correlations (0.24 – 1.00) and phenotypic correlations (0.08 – 0.39). Both sets of correlations follow the expected biological trend of decreasing as the distance between ages increases. The PE effect accounted for 2.31 kg<sup>2</sup> out of a total variance of 14.03 kg<sup>2</sup> in the model with a 0.5 correlation, 2.27 kg<sup>2</sup> out of a total variance of 14.15 kg<sup>2</sup> in the 0.1 correlation, and 2.34 kg<sup>2</sup> out of 14.05 kg<sup>2</sup> in the 0.9 correlation model. This suggests that the permanent environment effect was largely similar across the three applied correlations and the constraining of the correlation did not impact the permanent environment effect.

The results from the SBF model with a correlation of 0.5 are displayed in Table 3.5 These results also show a heritability that ranges from 0.33 to 0.49 across an age range of 1 to 376 days of age. This range of heritability values is similar to published examples for sheep growth that range between 0.10 and 0.40. The table includes genetic (0.34 – 1.00) and phenotypic (0.14

– 0.47) correlations which also follow an expected biological trend of decreasing as the difference between age points increases.

Heritabilities increased with age to reasonably high estimates at 376 days for SBF and 197 days for RC. In general, genetic correlations were high between close ages and much lower between ages further apart

**Table 3.4:** Phenotypic model summary values.

	<b>Polynomial Regression R<sup>2</sup> values</b>	<b>Spline</b>	<b>Gompertz:</b>
<b>Ramcompare</b>	Order 2: 0.826 Order 3: 0.835 Order 4: 0.837 Order 5: 0.837 Order 6: 0.838	R <sup>2</sup> : 0.615 P <.001	Residual standard error: 5.344 on 15053 df  Equation: $y = 43.12 e^{(-1.73 * 0.98^t)}$
<b>Scottish Blackface</b>	Order 2: 0.870 Order 3: 0.872 Order 4: 0.877 Order 5: 0.877 Order 6: 0.879	R <sup>2</sup> : 0.942 p < 0.001	Residual standard error: 3.561 on 23253 df  Equation: $y = 45.52 e^{(2.19 * 0.98^t)}$

**Table 3.5:** Genetic (below diagonal) and phenotypic (above diagonal) correlations among several weights at different ages with heritabilities (bold) on the diagonal from the random regression model for RC.

	<b>1</b>	<b>30</b>	<b>60</b>	<b>120</b>	<b>170</b>	<b>190</b>	<b>197</b>
<b>1</b>	<b>0.30</b>	0.28	0.25	0.18	0.11	0.09	0.08
<b>30</b>	0.98	<b>0.24</b>	0.25	0.20	0.16	0.14	0.14
<b>60</b>	0.91	0.97	<b>0.24</b>	0.23	0.20	0.19	0.19
<b>120</b>	0.62	0.76	0.89	<b>0.27</b>	0.29	0.29	0.29
<b>170</b>	0.35	0.53	0.71	0.95	<b>0.34</b>	0.36	0.36
<b>191</b>	0.26	0.44	0.64	0.92	1.00	<b>0.38</b>	0.39
<b>197</b>	0.24	0.42	0.62	0.91	0.99	1.00	<b>0.39</b>

1. There were 32,397 records for age 1, 35 for 30, 1,832 for 60, 250 for 120, 25 for 170, 1 for 191 and 1 for 197.

**Table 3.6:** Genetic (below diagonal) and phenotypic (above diagonal) correlations among several ages in days with heritabilities (bold) on the diagonal from the random regression model for SBF Data.

	<b>1</b>	<b>30</b>	<b>60</b>	<b>120</b>	<b>180</b>	<b>200</b>	<b>280</b>	<b>330</b>	<b>376</b>
<b>1</b>	<b>0.37</b>	0.36	0.35	0.32	0.28	0.27	0.21	0.18	0.14
<b>30</b>	1.00	<b>0.36</b>	0.35	0.32	0.29	0.28	0.23	0.21	0.18
<b>60</b>	0.98	0.99	<b>0.34</b>	0.33	0.30	0.30	0.41	0.24	0.21
<b>120</b>	0.91	0.94	0.97	<b>0.33</b>	0.32	0.32	0.31	0.29	0.28
<b>180</b>	0.79	0.84	0.90	0.97	<b>0.34</b>	0.34	0.35	0.35	0.34
<b>200</b>	0.74	0.80	0.86	0.95	1.00	<b>0.35</b>	0.36	0.36	0.36
<b>280</b>	0.54	0.62	0.70	0.85	0.95	0.97	<b>0.40</b>	0.42	0.43
<b>330</b>	0.43	0.51	0.60	0.77	0.90	0.93	0.99	<b>0.45</b>	0.47
<b>376</b>	0.34	0.43	0.52	0.70	0.85	0.88	0.97	1.00	<b>0.49</b>

1. There were 5,153 records at 1, 11 at 30, 150 at 60, 191 at 120, 10 at 180, 20 at 200, 31 at 280, 2 at 330, and 3 at 376.

### 3.3.3 Relationship between Growth and Carcass Traits

#### 3.3.3.1 Selection Index

For the selection index calculations, the G, R and P matrices for the four weights were as follows:

**Table 3.7:** G Matrix calculated for the Scottish Blackface data from multivariate models.

G matrix	BW	WW	SW	MW
BW	0.24	0.26	0.33	0.23
WW	0.26	4.72	3.51	1.54
SW	0.33	3.51	9.41	2.48
MW	0.23	1.54	2.48	7.57

**Table 3.8:** The R matrix calculated for the Scottish Blackface data from multivariate models.

R matrix	BW	WW	SW	MW
BW	0.14	0.34	0.39	0.22
WW	0.34	5.21	5.50	2.19
SW	0.39	5.50	7.91	3.50
MW	0.22	2.19	3.50	7.58

**Table 3.9:** The P matrix calculate for the Scottish Blackface data from multivariate models.

R matrix	BW	WW	SW	MW
BW	0.38	0.61	0.72	0.46
WW	0.61	9.93	9.01	3.72
SW	0.72	9.01	17.32	5.99
MW	0.46	3.72	5.99	15.15

The results for the selection index including all four weights for the slaughter weight carcass trait are included in Table 3.10. The accuracy of the slaughter weight index was 0.68. The way the difference in accuracy was calculated, as a ratio, the closer the value is to 1 the smaller that weight's contribution is to the overall index. In this case, the fourth weight makes the largest contribution to the overall selection index. However, it should be noted that the date of slaughter suggests that the fourth weight and slaughter weight are very similar if not identical traits in most cases for this dataset. The predictive ability therefore of the fourth weight for the slaughter weight is minimal. The b values indicate the impact of improving each weight by one standard deviation. For example, in this case improving the first weight by one standard deviation would result in a difference in slaughter weight of magnitude -0.14 standard deviations. The other

weights would have a change of -0.01, -0.06 and 0.60 for second weight, third weight and fourth weight respectively.

**Table 3.10:** Selection index results for slaughter weight.

<b>Weights</b>	<b>b value<sup>1</sup></b>	<b>Accuracy without weight</b>	<b>Difference in Accuracy</b>
First Weight	-0.14554	0.6762	0.9993
Second Weight	-0.01112	0.6877	1.016
Third Weight	-0.05699	0.6748	0.9972
Fourth Weight	0.604219	0.1816	0.2683

1. The B value is the index coefficient by which the phenotypic measurements are multiplied in the selection index; indicating the impact of increasing each weight by one standard deviation

The results for the selection index for carcass weight are displayed in Table 3.11. The accuracy of the selection index is 0.69. The fourth weight makes the largest contribution to the index. Here increasing each weight would correspond to change of magnitude - 0.12, 0.07, -0.15 and 0.32 standard deviations for the first weight, second weight, third weight and fourth weight respectively.

**Table 3.11:** The selection index results for carcass weight

<b>Weights</b>	<b>b value<sup>1</sup></b>	<b>Accuracy without weight</b>	<b>Difference in Accuracy</b>
First Weight	-0.11881	0.6856	0.9982
Second Weight	0.068361	0.6808	0.9912
Third Weight	-0.14508	0.6393	0.9310
Fourth Weight	0.317923	0.1361	0.1982

1. The B value is the index coefficient by which the phenotypic measurements are multiplied in the selection index; indicating the impact of increasing each weight by one standard deviation.

The results for the selection index for conformation are displayed in Table 3.12. The accuracy of the selection index is 0.11. The third weight makes the largest contribution to the overall index. Here increasing each weight would correspond to a change in the conformation score of magnitude 0.03,  $-1.39 \times 10^{-4}$ ,  $-1.60 \times 10^{-4}$  and  $-1.70 \times 10^{-3}$  for the first weight, second weight, third weight and fourth weight respectively.

**Table 3.12:** The results for the selection index for conformation

<b>Weights</b>	<b>b value<sup>1</sup></b>	<b>Accuracy without weight</b>	<b>Difference in Accuracy</b>
First Weight	0.02811	0.0357	0.3110
Second Weight	-0.000139	0.0495	0.4621
Third Weight	-0.00016	0.0000	0.000
Fourth Weight	-0.00170	0.1072	0.9349

1. The B value is the index coefficient by which the phenotypic measurements are multiplied in the selection index; indicating the impact of increasing each weight by one standard deviation.

The results for the selection index for fat class are displayed in Table 3.13. The accuracy of the selection index is 0.09. The fourth weight makes the largest contribution to the overall index. Here increasing the weight by one standard deviation would correspond to a change in fat class of magnitude 0.01,  $-6.6 \times 10^{-3}$ ,  $2.6 \times 10^{-3}$  and  $-2.2 \times 10^{-2}$  for first weight, second weight, third weight and fourth weight respectively.

**Table 3.13:** The results for the selection index for fat class.

<b>Weights</b>	<b>b value<sup>1</sup></b>	<b>Accuracy without weight</b>	<b>Difference in Accuracy</b>
First Weight	0.012202895	0.0844	0.9594
Second Weight	-0.006614635	0.0873	0.9922
Third Weight	0.00259437	0.0798	0.9073
Fourth Weight	-0.022333451	0.0036	0.0412

1. The B value is the index coefficient by which the phenotypic measurements are multiplied in the selection index; indicating the impact of increasing each weight by one standard deviation.

### ***3.3.3.2 Random regression modelling of growth and carcass traits***

The bivariate random regression models converged for two (fat class and conformation) of the four possible carcass traits. Tables 3.13 and 3.14 included correlations between

the weights as well as between each weight and a carcass trait. The heritability values for the weights are higher in these models that include carcass traits compared to the growth models without any carcass traits. These values appear to be inflated compared to the literature estimates which are more similar to the previous growth models. Potential explanations will be taken up in the discussion section of this study. The genetic correlation between conformation and body weight (0.17-0.84) was similar to the genetic correlation between fat class and body weight (0.31-0.85). Both correlations for conformation and fat class increase as the age at which the body weight was measured increased. This is consistent with biological expectation as the impact of body weight on conformation as well as fat class would be more pronounced as the animal ages. However, there seems to be little overall phenotypic relationship between body weight and conformation in this dataset which may be a consequence of the data structure as the uncaptured environmental variation would have to be large for the phenotypic correlation to be substantially smaller than the genetic correlation. It could also be impacted by variability in management of the animals wherein animals with good growth rates are slaughtered off grass while animals who are slower growing are managed in housing with hard feed. Fat class does show a stronger phenotypic correlation with body weight and follows the expected trend of increasing as the animal ages. Now it should be noted that the random regression model is able to model this relationship between fat class and growth in different animals as those animals with good growth rates are slaughtered earlier in life and therefore do not have weights at the high end of the age distribution. The model therefore is estimating a relationship between the two traits but not necessarily those two traits in the same individual animal. When the results from the random regression models are compared to the selection index for these two traits, the importance of body weight at higher ages is consistent.

For both conformation and fat class, the fourth weight makes the largest contribution to the accuracy of the overall selection index. However, the accuracy of the selection index for conformation is low compared to the selection index for slaughter weight or carcass weight which may be due to the data structure which could also account for the low phenotypic correlations between carcass trait and body weight in the random regression model

**Table 3.14:** A table of heritabilities down the diagonal with genetic correlations below the diagonal and phenotypic correlations above the diagonal for weights throughout an age distribution of 1 to 376 days and Conformation.

	<b>1</b>	<b>30</b>	<b>70</b>	<b>83</b>	<b>100</b>	<b>160</b>	<b>250</b>	<b>310</b>	<b>343</b>	<b>376</b>	<b>Carcass</b>
<b>1</b>	<b>0.67</b>	0.64	0.58	0.56	0.52	0.34	0.02	-0.16	-0.23	-0.30	0.00
<b>30</b>	1.00	<b>0.62</b>	0.57	0.55	0.52	0.36	0.06	-0.10	-0.17	-0.23	0.00
<b>70</b>	0.97	0.99	<b>0.54</b>	0.52	0.50	0.38	0.14	0.00	-0.06	-0.11	0.00
<b>83</b>	0.95	0.97	1.00	<b>0.51</b>	0.49	0.39	0.17	0.04	-0.02	-0.07	0.00
<b>100</b>	0.91	0.95	0.99	1.00	<b>0.48</b>	0.40	0.21	0.09	0.03	-0.01	0.00
<b>160</b>	0.81	0.72	0.82	0.86	0.90	<b>0.40</b>	0.34	0.28	0.25	0.22	0.00
<b>250</b>	0.03	0.12	0.28	0.34	0.43	0.78	<b>0.49</b>	0.52	0.53	0.53	0.01
<b>310</b>	-0.25	-0.16	0.00	0.07	0.16	0.57	0.96	<b>0.61</b>	0.63	0.65	0.01
<b>343</b>	-0.35	-0.26	-0.10	-0.04	0.06	0.48	0.93	0.99	<b>0.67</b>	0.69	0.01
<b>376</b>	-0.42	-0.34	-0.18	-0.12	-0.02	0.41	0.89	0.98	0.76	<b>0.72</b>	0.01
<b>Carcass</b>	0.17	0.24	0.35	0.40	0.46	0.67	0.74	0.67	0.63	0.84	<b>0.31</b>

**Table 3.15:** A table of heritabilities down the diagonal with genetic correlations below the diagonal and phenotypic correlations above the diagonal for weights throughout an age distribution of 1 to 376 days and Fat Class

	<b>1</b>	<b>30</b>	<b>70</b>	<b>83</b>	<b>100</b>	<b>160</b>	<b>250</b>	<b>310</b>	<b>343</b>	<b>376</b>	<b>Carcass</b>
<b>30</b>	0.99	<b>0.42</b>	0.36	0.34	0.31	0.17	-0.02	-0.11	-0.14	-0.18	0.25
<b>70</b>	0.99	0.96	<b>0.34</b>	0.33	0.31	0.19	0.10	0.03	0.00	-0.03	0.25
<b>83</b>	0.95	0.93	0.99	<b>0.32</b>	0.31	0.25	0.14	0.07	0.05	0.03	0.28
<b>100</b>	0.78	0.86	0.97	0.99	<b>0.30</b>	0.27	0.19	0.14	0.12	0.10	0.31
<b>130</b>	0.58	0.69	0.86	0.91	0.96	0.30	0.28	0.25	0.24	0.22	0.37
<b>160</b>	0.34	0.47	0.73	0.76	0.85	<b>0.33</b>	0.36	0.35	0.35	0.34	0.42
<b>250</b>	-0.17	-0.03	0.23	0.33	0.47	0.87	<b>0.52</b>	0.57	0.57	0.60	0.50
<b>310</b>	-0.34	-0.20	0.06	0.12	0.32	0.77	0.99	<b>0.65</b>	0.67	0.69	0.51
<b>343</b>	-0.40	-0.27	0.00	0.11	0.25	0.73	0.98	1.00	<b>0.70</b>	0.72	0.51
<b>376</b>	-0.44	-0.31	-0.05	0.06	0.21	0.69	0.96	0.99	1.00	<b>0.75</b>	0.51
<b>Carcass</b>	0.49	0.31	0.40	0.56	0.38	0.46	0.68	0.86	0.96	0.85	<b>0.48</b>

## **3.4 DISCUSSION**

### ***3.4.1 Phenotypic Curves***

The curve from the phenotypic analyses indicates the feasibility of fitting polynomial regressions of several orders. The polynomial regressions in Excel increased in R squared value as the order of the polynomial regression increased – a trend that must be balanced with the increasing degrees of freedom in the model. The spline functions had reasonably high R squared values with SBF having a higher R squared value indicating a better goodness of fit. However, analyses at the genetic level with such a high number of knots will not be feasible. The Gompertz models produced similar model equations for both datasets. The phenotypic modelling was used to understand how the phenotypic data was distributed. Given that all phenotypic models produced some form of curve it was difficult to determine which if any of them were clearly representative of the phenotypic distribution of the data. In addition, Gompertz may not be easy to fit in a random regression setting.

### ***3.4.2 Genotypic analysis***

The application of random regression models to commercial data is still being investigated. This study incorporated commercial data and research data to compare the application in both contexts. Despite the different environments, both the commercial and research flocks had at most four records for weight. The minimum number of records required for random regression modelling is still a question to be answered. In this study, the total number of records available appears to be too small resulting in high correlations between the slope and intercept of the polynomial. A lack of records per animal and a lack of records for parents has been noted to

impact genetic modelling and limit the accuracy of genetic estimates (Maniatis and Pollott, 2003). Additionally, these growth models did not include maternal effects or the covariance between the direct and maternal effects likely due to the same lack of records for dams, which may limit the accuracy of the genetic estimates (Maniatis and Pollott, 2002a, 2002b; Safari, Gilmour and Fogarty, 2005; Conington et al., 1995).

Setting the correlations to predefined values provided more reasonable heritability estimates that were more consistent with the average values available for growth across published literature (Amer et al., 2007; Lambe et al., 2008). Although the heritability for birth weight did not follow the general pattern of increasing over time, this is reasonable if birth weight is under different genetic control than growth at other points during a lamb's life. The random regression model assumes that each weight is correlated to but genetically distinct from the others, meaning a genetic correlation of unity is not expected in these models. These models follow the expectation except between ages close together. It should also be noted that standard error estimates were not possible to compute for these models due to the issues with convergence and subsequent constraining of the correlation making estimation of standard error computationally demanding. It was therefore not possible with the available computational software to estimate these values. This makes it difficult to ascertain if the estimates provided are robust and representative of the true genetic value for these traits. Additionally, it is difficult to compare these heritability estimates against other studies because the reliability of these estimates is unknown compared to those where standard error, indicating how reliably the same parameters would be produced if the sample size was evaluated again, is presented.

These results provide an initial basis to investigate further the impact of sparse data on the performance of random regression models, suggesting that sparse data yield very high correlation estimates which can be overcome by fixing the correlation to a plausible value. It indicates that further investigation is needed into higher order Legendre Polynomials with fixed correlations and more flexible functions such as the Spline could be helpful in overcoming high correlations produced by these models. Additionally, this study provides additional evidence that one of the major qualifications for using a random regression model in growth is having many records that are evenly distributed throughout the age distribution. This implies use of random regression models in genetic evaluations require data collectors to have a strategy to record across the trajectory being modelled.

### ***3.4.3 Relationship between Growth and Carcass Traits***

#### ***3.4.3.1 Selection Index***

The discussion of the results for the selection index will centre on the three questions that this part of the study sought to answer for each carcass trait we investigated. For slaughter weight, the accuracy of the selection index was 0.6767 which means that we can predict slaughter weight from body weight with an accuracy of 67%. The second investigation, into which weight is the most important or contributes the most to the index, is directed towards the relative accuracies of the selection indices where one weight has been removed. For slaughter weight, the accuracy values suggest that the fourth weight plays the largest part in the accurate selection of slaughter weight. The low accuracy of 0.18 for the selection index without mature weight is evidence of this conclusion. Additionally, the ratio between the accuracy of the selection index and the selection index without mature weight is furthest from one.

The final question for this portion of the study was what the impact would be on the carcass trait if the value of the weight traits were increased by 1 standard deviation. The b values provide the answer for this question. For slaughter weight, an improvement by 1 standard deviation in the first weight would correspond to a difference in slaughter weight of magnitude -0.14 standard deviations. For the second weight there would be a change of magnitude -0.01 standard deviations. An improvement in the third weight would result in a change of magnitude -0.06 standard deviations. For the fourth weight there would be a change of magnitude 0.60 standard deviations. Thus, for slaughter weight the strongest point of positive genetic relationship was at the fourth weight. This result is expected because the fourth weight is closest to the slaughter weight, and in many cases is an identical weight for the lambs included in the dataset. There is a strong relationship between fourth weight and slaughter weight. For producers, focusing on increasing the fourth weight by 1 unit, would result in an increase of 0.60 standard deviations from the average slaughter weight. It should be noted that large weights are not as much a producer's goal as a weight that falls within the specification for their particular breed of sheep. Thus, if a producer needs to maintain or lower a weight in order to meet the specification the largest opportunity for intervention is in the fourth weight as well. Bringing this down will correspond to a lower slaughter weight that is lowered by 0.60 standard deviations from the average slaughter weight value for the given population. Additionally, the results suggest that the fourth weight or more mature weight is a different objective for terminal and maternal breeders (McHugh et al., 2022).

For the other weights, the relationship between them and the slaughter weight is weaker as they are not as significant determinants of the slaughter weight's accuracy. Thus, it is difficult to form any practical recommendation for the other weights in the selection index because any change in them is likely to have a minute impact on the carcass trait. The selection index does suggest that an increase in the first weight would change the slaughter weight by -0.14 standard deviations from the trait mean, for the second weight it would increase the slaughter weight by -0.01, and for the third weight an increase would correspond to a change of -0.06 standard deviations from the trait mean. Again, it must be stressed that the relationship between the carcass trait and these three weights is not strong. Thus, a change in any of these weights would have a minute, if any, impact on the carcass trait. Thus, the producer should focus mostly on the weight that has the largest impact on the carcass trait. However, the applicability of this finding depends heavily on the management of the sheep. In this case the sheep are not reared to a certain age but to a target live weight and condition score that aims to correspond to carcass weight requirements and EUROP grade criteria. This means that a lamb that grew quickly to attain its target live weight by weaning could have the same carcass weight as a lamb that grew more slowly only achieving its target live weight at the third weighing. Thus the selection index may be unable to provide a particularly helpful management goal for this system which focuses instead on achieving a target weight quickly rather than focusing on achieving that weight at a specific weighing point.

The selection index for carcass weight shows a similar accuracy to slaughter weight. Carcass weight can be predicted from body weight with an accuracy of 0.68. Here again, the fourth weight makes the largest contribution to the selection index. When it is removed the selection index value is most different from its original value. The

relative contribution of the other body weight traits if ranked from highest to lowest is, third, second and first weight. Finally, if the first weight was improved by 1 standard deviation, then the carcass weight would change by -0.12 standard deviations. For the second weight there would be a change of 0.07 standard deviations. An improvement in the third weight would correspond to a change of -0.15 standard deviations. For the fourth weight there would be a change of 0.32 standard deviations. Again, the strongest genetic relationship between the carcass trait and body weight is present at the fourth weight. Thus, the largest opportunity a producer must achieve their goal carcass weight is in the fourth weight. If a producer wants a higher carcass weight, increasing the fourth weight will have the largest positive impact of 0.32 standard deviations from the average carcass weight. If a producer wants a lower carcass weight, decreasing the fourth weight will have the largest negative impact of the same magnitude. Again, the impact of other weights. Here again the management program may impact the degree to which these findings are helpful as the carcass weight may be the same for two animals who are slaughtered at different ages depending on their growth rate. Overall there does seem to be a relationship between body live weight at these four points and the carcass traits. Indeed, selection indices have been successfully used in the lean sire index for example to achieve carcass characteristics in lambs by selecting on particular live weights and carcass characteristics in their sires (Marquez et al., 2013).

The accuracy of the selection index for conformation is much less than the previous two carcass traits at 0.11. Conformation can be predicted from body weight with an accuracy of 11%. The largest contribution to the selection index is made by the third weight based on the ratio difference between the original selection index and the

selection index without the third weight being furthest from 1. Following from the third weight, the largest contribution is made by the first followed by the second then the fourth. Finally, an improvement in body weight by 1 standard deviation would result in a change to conformation score of 0.03 standard deviations for the first weight. For the second weight the change would be of magnitude -0.01 standard deviations. For the third weight the change would be of magnitude -0.01 standard deviations. An improvement in the fourth weight would correspond to a change of -0.01 standard deviations. Here it is notable that the contribution of the third weight does not accord with the magnitude of the b value. Thus, it is difficult to establish a practical recommendation from the selection index for conformation. The selection index is largely controlled by the third weight but the impact of increasing the body weight has a minute negative impact on the conformation. This may be due to the data structure and the nature of the conformation variable which is a category rather than a numeric number.

Finally, the accuracy of the selection index for fat class is 0.08 meaning fat class can be predicted from body weight with an accuracy of 8%. The largest contribution to the index is made by the fourth weight with a difference of 0.04 which is the furthest from a value of 1 when compared to the differences for the other weights. The contribution from the other weights ranked from highest to lowest would be third then first weight and second weight. Finally, the change in fat class corresponding to an improvement of 1 standard deviation in first weight would be -0.02 standard deviations. For the second weight the change in fat class would be of magnitude 0.01 standard deviations. For the third weight, the change in fat class would be of magnitude -0.01 standard deviations. For the fourth weight, an improvement of 1 standard deviation would correspond to a

change in fat class of magnitude 0.01 standard deviations. For fat class, the fourth weight makes the largest contribution to the overall selection index suggesting that it has the largest impact on fat class. Improving the fourth weight has an improvement of 0.01 standard deviations above the mean from the average fat class value. Again, it is important to remember that fat class is a categorical variable.

Overall, the prediction of slaughter weight and carcass weight from body weight can be performed with reasonable accuracy. Conformation and fat class are much less accurately predicted from body weight. In this case, slaughter weight and the final weight can be considered identical in most cases. However, the prediction of carcass weight from body weight is consistent with other studies that sought to predict carcass traits from live weights (Connington et al., 1995; Lambe et al., 2008). The less clear relationship between live weights and fat class or conformation could be due to a lack of granularity between live body weight and these two carcass traits. Lambe et al. (2008) established a genetic correlation between fat measured by computer tomography and ultrasound back fat depth at weaning and subjective fatness grade at slaughter. Additionally, the same study established ultrasound muscle depth at weaning and subjective conformation score at slaughter had high genetic correlations with the muscularity indices used in the study that focused on the hind leg and lumbar regions (Lambe et al., 2008). Therefore, it could be that live weight was not representative of the components in live weight, including muscle, fat cover, or bone density, that are genetically correlated to these two carcass traits.

The contributions of body weights are not equal for these carcass trait selection indices. For the slaughter weight and carcass weight an improvement in mature weight

results in the largest improvement in the carcass trait. For the conformation the most important weight appears to be the third weight from the differences in accuracy between the original selection index and the index without third weight. However, the b values contradict this suggesting that a change in the birth weight would have the largest impact on the conformation of the animal. For fat class the prediction accuracy is the lowest. The largest contribution to the selection index is made by the fourth weight once again and this corresponds to the b values as well suggesting an improvement in the fourth weight would have the largest magnitude improvement in fat class as well.

#### ***3.4.3.2 Random Regression Analysis of Carcass Traits in SBF***

The bivariate random regression model converged only for fat class and conformation. The lack of convergence for slaughter weight in particular may be impacted by the data structure as it was subsequently determined that the slaughter weights and the fourth weights were not always collected at the same ages, making it impossible for the model to distinguish between them. The results from the fat class and conformation models showed higher heritabilities for the body weight trait in both models. The inflated heritabilities may be due to the inclusion of the carcass trait in the models or a lack of body weight records throughout the entire age distribution. When the results from the carcass models are compared with the models without the carcass trait, the heritabilities with the carcass traits are higher at (0.30 – 0.75) for fat class and (0.40 -0.72) for conformation score compared to (0.33-0.49). The low end of the ranges are similar but the carcass models have a much higher maximum heritability. The heritability increase may be caused by the increase in genetic variance and reduction in residual variance

between the model with only weight compared to the models that included a carcass trait. The genetic variance in the weight model accounted for 48% of the overall variance while it accounted for 55% and 57.5% in the models with fat class and conformation respectively. The residual variance accounted for 44% of the variance for the model with only weight and 35% for both the carcass trait models. The permanent environmental effect stayed largely the same between all the models. It accounted for 8%, 10% and 7.5% of the overall variance in the weight model, the fat class model and the conformation model respectively. The addition of the carcass trait increased genetic variance and reduced residual variance because the relationships between the carcass trait and body weight meant additional information contributed to the estimates for body weights.

It was not possible to compute standard errors for the heritability estimates meaning it is not possible to compare the estimates with or without carcass traits because the two different carcass traits in the random regression model result in different variance parameters in the models meaning they cannot be compared directly with one another or with the growth model. The standard errors were also inestimable due to the computational demand introduced by the convergence issues with the models prior to fixing the correlations between the intercept and slope of the Legendre polynomial. It should therefore be noted that there is not a consistent and comparable measure of the reliability of these genetic parameters. However, the inclusion of the carcass traits and the data structure may have introduced some modelling difficulties due to data imbalance given the data structure. The weights model went through 7 iterations before converging, the conformation model went through 12 iterations, and the fat class models went through 7 iterations. The higher number of iterations for the conformation

model could also contribute to the higher heritability values because a higher number of iterations can indicate a greater degree of statistical challenge in accurately estimating variance figures.

Additionally, although each animal had four weight records, there is a significant reduction in record numbers near the end of the age distribution as most of the weight records are taken before 180 days of age. These results seem to be impacted significantly by the data structure. The heritability value for fat class at 0.48 is also inflated compared to previous studies (Maxa et al., 2007; Massender et al., 2019). Texel sheep fat class has a heritability of 0.11 and Shropshire has a heritability of 0.19 (Maxa et al., 2007). However, the heritability of the conformation at 0.31 was more consistent with other published estimates of the same trait which are around 0.1 – 0.4 (Massender et al., 2019). The random regression model for conformation shows a genetic correlation between body weight and conformation that increases as age increases. The correlation is considerable, reaching 0.84 at the top of the age range. This suggests that there is a strong genetic relationship between body weight and conformation. Although this is potentially consistent with other studies on the relationship between conformation and certain body weight components like muscularity, the genetic correlation is higher than that reported in other studies (Lambe et al., 2007; Lambe et al., 2008). This could be due to the data structure or the absence of a maternal effect fitted into the model (Maniatis and Pollott, 2002c, 2003).

However, it should be noted that the phenotypic correlation remains small throughout the age distribution. This could be due to the data structure or some uncaptured environmental variance. The results for the fat class show a similar genetic trend

indicating a substantial genetic relationship between fat class and body weight with the strongest relationship at the top of the age range. Fat class shows a consistent phenotypic correlation suggesting that there is a phenotypic relationship between body weight and fat class that increases as the age increases. This is expected if body weight has an impact on fat deposition which would be consistent with Lambe et al., 2008 finding that there was a positive genetic correlation between computer tomography fat and subjective fat class at slaughter. These results are more consistent than those from the selection index as they have clear genetic correlations that increase over time while the selection index accuracy was low and the magnitude of the effect of improving a given body weight did not correspond with the contribution that body weight made to the overall selection index. The random regression models offer a clear practical suggestion to focus on improving the body weight at the top of the age distribution to have the largest effect on conformation and fat class.

The comparison of these two methods suggests that the limitations of the dataset with a structure that is not normally distributed across the age range requires the use of a selection index for a more accurate treatment of the carcass traits. The bivariate random regression model inflated the heritability for the weight traits and for fat class. These inflated values are less accurate than those estimated from the original genetic models and inconsistent with published estimates of the same trait (Conington et al., 1995; Amer et al., 2007). However, in this case the combination of findings from the selection index and the random regression models offers a practical suggestion for all four carcass traits. The results from the selection index were most accurate for slaughter weight and carcass weight. According to those results, producers should focus on optimizing the fourth weight to have the largest impact on the slaughter weight and the carcass weight.

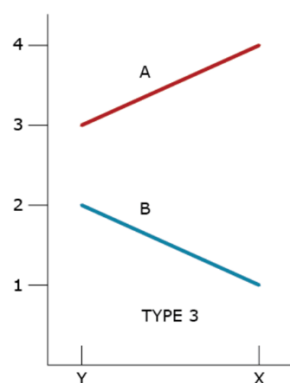
The random regression models converged for the two traits with the least accurate selection index: fat class and conformation. In both cases the random regression models showed a genetic relationship between body weight and conformation or fat class. This genetic relationship increased as the age increased with the highest correlation present at an age of 376 days. Thus, a producer can have the largest effect on fat class and conformation by focusing on the weight taken at the highest age.

Overall, this chapter has not only compared commercial and research data in random regression modelling but has also attempted to apply the random regression model to more comprehensive traits than just weight. The study found that random regression modelling is highly sensitive not only to the number of records per animal but also to the distribution of those records throughout the age distribution. This chapter has found that these models would have improved accuracy if there were more weight records distributed more evenly for both the SBF and RC datasets. Further research should investigate the minimum number of records required to achieve similar accuracy from random regression modelling techniques compared to those achieved through standard multi trait modeling.

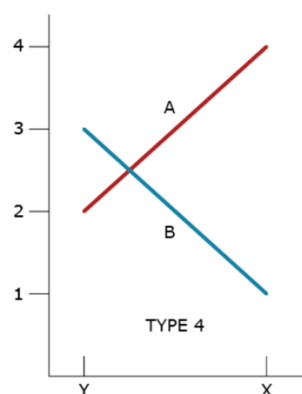
## CHAPTER FOUR: MODELLING GXE INTERACTION IN TERMINAL SIRES IN UK SHEEP POPULATIONS

### 4.1 INTRODUCTION

The fourth and final aim of this thesis was to assess whether there are genotype x environment (GxE) interactions in terminal sire breeds across the United Kingdom. These effects can alter the performance of specific sires in different environments across the UK which could mean that one sire may be the best choice for one environment but then perform poorly elsewhere; suggesting that a breeder should prefer a different sire whose progeny are more properly adapted for that environment. GxE interactions can impact an animal's performance and may decrease the accuracy of an EBV if they are not considered (Woolaston, 1987). The effects may cause scaling as displayed in Figure 4.1 wherein the ranking of animals' genetic potential may stay the same but the difference in performance may change in different environments. Or the GxE interaction may cause re-ranking as displayed in Figure 4.2 where a genotype that is the best performer in one environment is not the best performer in another environment (Falconer and Mackay, 1996).



**Figure 4.1:**Genotype response with GxE causing scaling but no reranking of genotypes.



**Figure 4.2:** Genotype response with GxE causing reranking of genotypes.

The overall ranking of animals may change if multiple traits in a selection index have scaling differences. These effects can inhibit genetic gain, animal performance, and the overall profitability of commercial sheep systems. Accounting for these interactions, where present, is essential to accurately predict the value of a particular genotype in a specific environment.

There are two main ways of identifying GxE interactions: the multi trait model and the reaction norm model. The multi trait model is most applicable in scenarios where environments can be completely distinguished from one another. For example, if a specific trait is being investigated in separate geographical regions. Jakobsen et al. (2009) investigated GxE in dairy sires across countries for protein yield, female fertility and longevity using multi trait models. Distinct environments can also be categorized using production system (Wallenbeck et al., 2009; Kelly et al., 2007) or feeding regime (Boettcher et al., 2003; Kearney et al., 2004a, 2004b). On the other hand, reaction norm models are generally used when the environment varies on a continuous scale. This means the environments are classed as more or less favourable for a trait; or the trait is more or less resilient. In this case resilience refers to the trait's consistency across environments. If a trait performs similarly across environments under different conditions, then it is considered more resilient. On the other hand, a trait that is impacted by

the environment is considered less resilient. Examples of environments that are continuous in scale include those affecting production levels like milk yield in a dairy herd and fertility descriptors (Kolmodin et al., 2004; Huquet et al., 2012; Zhang et al., 2019). Other studies incorporate continuous scales based on regional averages for temperature or other climatic conditions (McLaren et al., 2015). The majority of GxE studies are conducted in the dairy cattle industry. Although there are comparatively fewer GxE studies investigating commercial sheep systems, investigation of GxE in commercial sheep traits has been done in the UK (Maniatis and Pollott., 2002a; Pollott and Greeff, 2004). This is a growing area of interest as the international exchange of sheep genetics becomes more common. There are also concerns about the presence of GxE across various systems within a country as well as similar systems in different countries.

The impact of different geographical regions and different nutrition profiles on carcass traits and growth have been investigated (Bishop et al., 1996; Lewis et al., 2004; Macfarlane et al., 2004a, 2004b). Bishop et al., (1996) found that the rearing environment of Scottish Blackface lambs, in hill pasture or improved pasture, had a significant impact on backfat thickness. In that study, lambs reared in improved pastures had twice the backfat thickness of lambs from hill pastures. Lewis et al., (2004) investigated the impact of a nutritionally non-limiting food compared to that of a bulky food on lamb growth and carcass composition in Suffolk sheep, Scottish Blackface sheep and their reciprocal crosses. The study found that on the nutritionally non-limiting food, Suffolk sheep were less fat and more lean than both the Scottish Blackface and reciprocal crosses. The bulky food caused sheep across breeds, sexes and stages of maturity to have 0.810 as much fat and 1.063 as much lean compared to the other food. Macfarlane et al., (2004b) investigated the impact of different grazing diets in Suffolk sheep, Scottish Blackface, and their crosses on carcass composition and growth. The study found that Suffolk

lambs had higher growth rates than Scottish Blackface lambs on clover but not on ryegrass or a mixture of the two. There was no significant difference between Suffolk and crossbred sheep. Suffolk sheep also showed a larger reduction in growth rate when the nutritional environment became worse. Suffolk sheep were therefore more environmentally sensitive than Scottish Blackface sheep. However, Macfarlane et al., (2004a) found no effect of diet on carcass composition. Some studies have found significant impact of weather conditions on lamb weaning weights (Steinheim et al., 2004, 2008).

In this chapter, the significance of GxE interactions for commercial sheep production is examined and is important given that the data were collected as a part of AHDB's RamCompare project. This is the UK's national progeny test which has been running from 2015. The project uses nominated performance recorded rams from various terminal sire breeds across the UK. These rams have progeny across a large proportion of commercial flocks meaning that the genetic assessments this project produces can be used widely across the industry. This dataset is used to provide commercially appropriate EBVs for carcass weight, conformation, fat class, days to slaughter, and an overall carcass merit index. Thus, the RamCompare project is actively providing up to date genetic evaluations that seek to improve efficiency in commercial sheep production across a range of environmental conditions and management practices. Given this range of environments, there is potential for GxE effects to be present. If these are not accounted for then this widely used commercial genetic evaluation project may lack some accuracy. In this chapter, we investigated the presence of GxE effects in this progeny testing project that has an impact on the entire UK sheep industry.

Additionally, this study contributes to the growing body of research on GxE effects in commercial sheep which can be more difficult to model compared to experimental or controlled

environments. This is because GxE models rely on the identification of environmental factors. In commercial flocks, these factors are not as controlled as they can be in flocks which are maintained for research. There can be a range of factors that differentiate one farm environment from the next. There are therefore multiple ways of differentiating these farm environments. This can include climatic characteristics such as rain fall or average temperature (McLaren et al., 2015), the use of phenotypic means or deviations where contemporary groups or flocks are differentiated by the degree to which they deviate from the whole population average for a specific trait (Fikse et al., 2003; Ravagnolo et al., 2002), or the identification of contemporary groups where a flock effect or contemporary group effect is used to account for different environments (Pollott and Greeff, 2004).

In this study the GxE effect was investigated using two methodologies. In the first case a contemporary group effect was analysed using the flock as the contemporary group and representation of the environment. In this methodology a sire model was used with a flock interaction to assess how an individual sire's EBV changed across the flocks in which they had progeny. The second methodology employed the regression technique that has been used throughout the studies included in the thesis. The regression technique used a reaction norm model with phenotypic deviations, to identify GxE effects in commercial sheep flocks. Using two methodologies for this dataset contributes not only to the growing body of research on GxE interaction in sheep but also addresses the applicability of GxE modelling to commercial data. This study provides useful practical recommendations on the inclusion of GxE effects and their impact on the genetic evaluations for the UK's commercial sheep industry.

## **4.2 MATERIALS AND METHODS**

### ***4.2.1 Data Collection and Experimental Design***

The dataset for analysis was constructed to compare the performance of multiple terminal sires in flocks across the UK. Additional variables were included in the RamCompare (RC) dataset which was used in the previous chapter for random regression modelling for growth. These additional variables captured differences in the animal's environments. The dataset for GxE analysis included postcodes (12 levels), average temperature in Celsius (7.70 -11.40), and average rainfall in mm (662.6 -1559). A flock effect was then calculated as each flock's average deviation from the population average for the birth weight, weaning weight, scan weight, ultrasound muscle depth and ultrasound fat depth. There were 393 sires that were used in at least two of 10 total flocks which were incorporated into the GxE analysis. The number of flocks included was reduced from the RC dataset to include only those flocks with lambs sired from one of the 393 sires that had progeny in at least two flocks.

The breeding program for the Ramcompare project is significant for GxE analysis as it makes it possible to compare sire performance. The terminal sires being compared in this study all had progeny present in one flock. Then unique subsets of the sires would be used in the other flocks. This means that for most of the sires, the GxE effect, if present, would be observed only between the two flocks in which their progeny was present.

### ***4.2.2 Sire and Flock Interaction***

In the first method, using contemporary groups as the environmental indicator, the first step of the modelling was to construct a sire model to assess the significance of the sire effect for the

traits of interest. The sire model also allowed us to provide an initial EBV for each sire prior to the inclusion of the potential GxE effects. A sire model was constructed for each of the traits of interest which included birth weight, weaning weight, ultrasound muscle depth and ultrasound fat depth

The model was constructed as follows:

$$Y_{ktij} = Xb + Zs + e_{nktij} \quad (4.1)$$

Here  $y$  is the vector of observations for the traits of interest.  $Xb$  represents the fixed effects including sex (3 levels, Male, Female and Castrate) and year (4 levels). Other fixed effects like litter size and dam age were not available in the dataset. Age was not included in the model because it resulted in the models failing to converge. The  $Zs$  represents the random genetic effect of the sires and  $e$  is the residual error which is not captured by the model. Each of the models for the five traits of interest did converge and provide results. The significance of the sire effect was analyzed by performing a Log Likelihood ratio test between models with and without the sire effect.

Following on from the sire models, an environmental interaction was included in the model to assess how sire performance differed across the flocks. There was a random sire effect, a random flock effect and a random sire by flock interaction included in the sire model described by Model 4.1. This assessment allows GxE effects to be investigated through reranking of EBVs across the flocks (Brown et al., 2009). If there is substantial reranking where the correlations between the ranks is less than 0.85 then GxE effects are present (Falconer and Mackay, 1996). EBVs were calculated for each sire within each flock and ranked from most positive to least positive. Correlation values were then calculated between each sire's ranking

across the flocks in which they had progeny. Plots of the ranks for sires across flocks were constructed in R studio using “ggplot” (Rstudio Team, 2012).

### **4.2.3 Reaction Norm Modelling**

The dataset was then analyzed using a reaction norm model to investigate GxE effects. This analysis followed the methodology proposed by Fikse et al. (2003) in dairy cattle. The four variables of interest were analyzed using reaction norms for sires which were modelled as a random regression of the trait of interest on an environmental descriptor.

The model was structured as:

$$Y_{nktij} = F_i + \text{sire}(env) + e_{nktij} \quad (4.2)$$

The  $\text{sire}(env)$  is a random regression on each flock’s average for the trait of interest.  $Y$  is the  $n$ th weight observation of the  $k$ th animal at time  $t$  belonging to the  $i$ th fixed factors and the  $j$ th group;  $F_i$  represents fixed effects consisting of Sex (3 levels), Year (4 levels), season of measurement of the trait of interest (2 levels for weaning weight, muscle depth and scan weight and 1 level for weaning weight), and a combination season year variable (4 levels for scan weight, muscle depth and fat depth and 7 levels for weaning weight) and  $e$  is the random residual variance which is not accounted for by the model. The season and season year variables were included to account for potential differences in the environmental conditions between the seasons and years where records were collected. The reaction norm model using the random regression methodology is considered to be potentially more able to account for these differences in environmental conditions that cannot be captured in fixed effects like

contemporary group compared to multi-trait models. The flock averages for each trait; birth weight, weaning weight, scan weight, and ultrasound muscle depth and ultrasound fat depth were the respective environmental descriptors in each of the reaction norm models that were constructed for each trait. This methodology was chosen because there was a lack of additional information that could have been used to create an environmental scale or group flocks together based on environmental criteria. The data suggested that the environments were largely similar as indicated by rainfall and temperature and no additional information was provided about the farming system or environment in which the farm was located. Thus, using a phenotypic indicator of environment was identified as the best alternative to using environmental grouping or scaling.

The random regression on the flock averages were modelled using Legendre polynomials. For birth weight, Legendre polynomials of order 1-3 were fitted without convergence, for weaning weight an order of 1 was fitted and converged, for scan weight orders of 1 through 3 were fitted with 1 converging, for muscle depth an order of 1 was fitted and converged, and for fat depth polynomials of 1-3 were modelled without convergence. Log likelihood values were used to determine the significance of the GxE effect using a likelihood ratio test. All genetic analyses were conducted in ASREML (Gilmour et al., 2016).

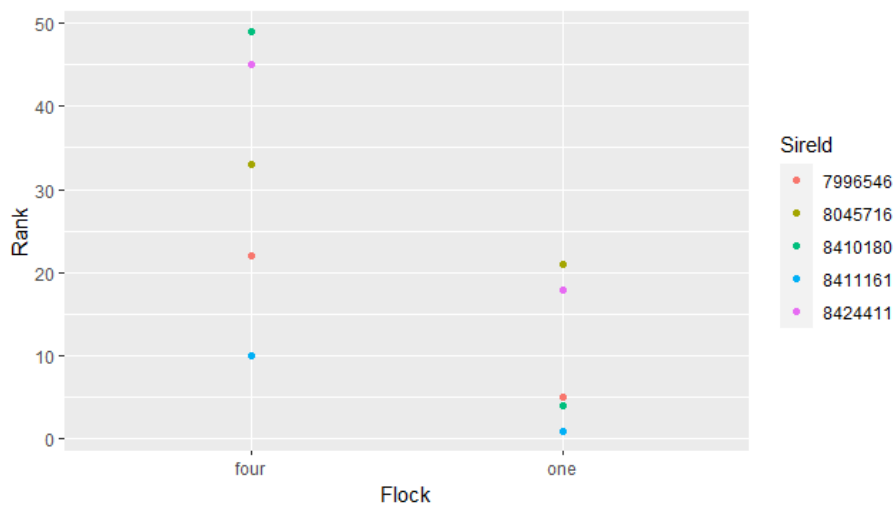
## **4.3 RESULTS**

### ***4.3.1 Sire and Flock Interaction***

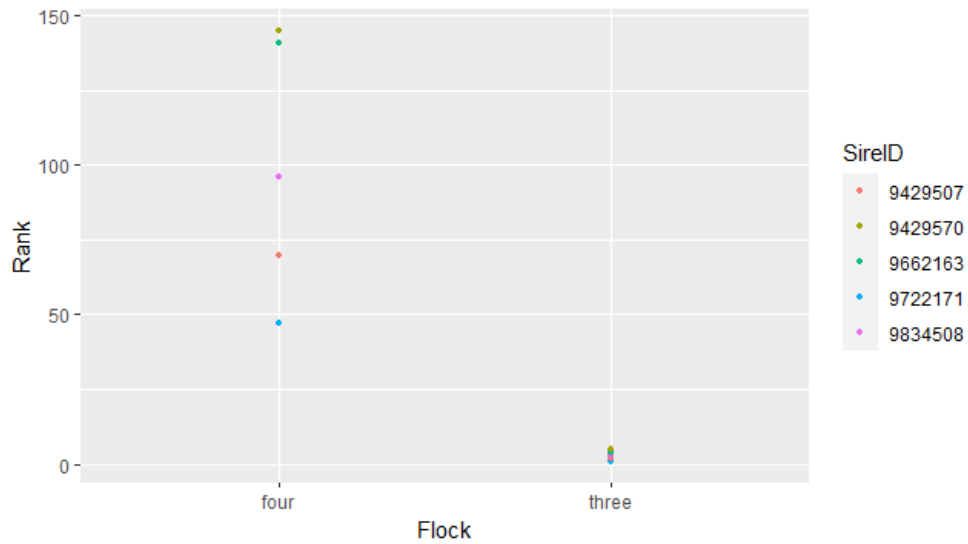
The first set of sire models described by Model 4.1 showed that the sire effect was a significant genetic effect for each of the traits of interest except for fat depth where the model could not achieve convergence. It was hypothesized that the model could not achieve convergence due

to missing records for this trait. When the EBVs of sires were calculated across flocks there was evidence of GxE effects in terms of reranking of sires for the other traits of interest. The correlation values between the ranks for each flock revealed reranking in some but not all the flocks suggesting that although there are GxE effects present for the traits, they are not causing significant reranking, considered in terms of the correlation of 0.85 or less, in all the flocks (Falconer and Mackay, 1996). For birth weight there were three out of ten flocks with reranking. For weaning weight there were two of ten flocks with reranking. For scan weight there were three out of ten flocks with reranking of sires. For ultrasound muscle depth there were eight of ten flocks with reranking.

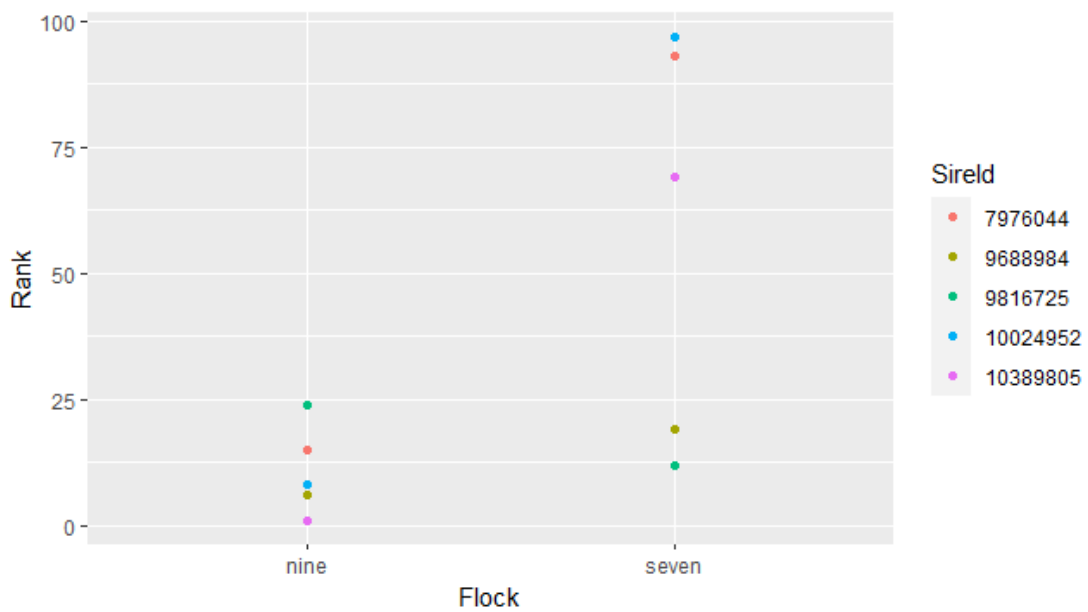
Plots for each trait were constructed showing five sires except in Plot 4.7 where an additional sire was shown due to the proximity of the sires EBV ranks, which made it difficult to visually identify the reranking with only five sires. One plot shows two flocks with reranking and one plot shows two flock without reranking.



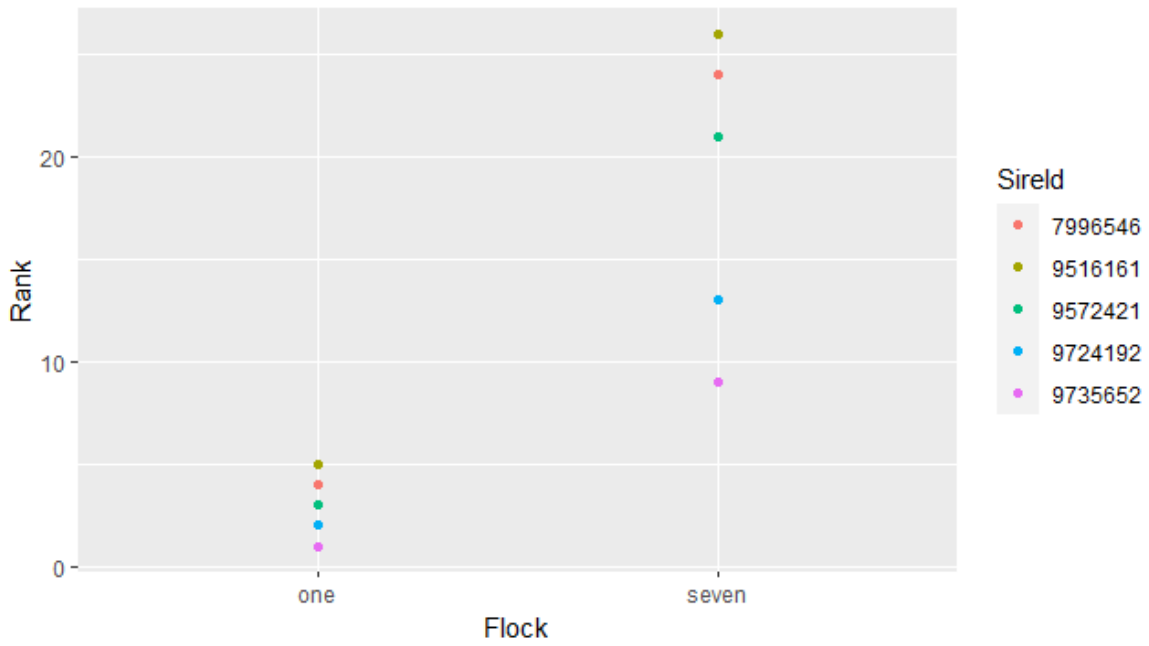
**Figure 4.3** Ranked EBVs for Muscle Depth for the same sires between two flocks showing re ranking.



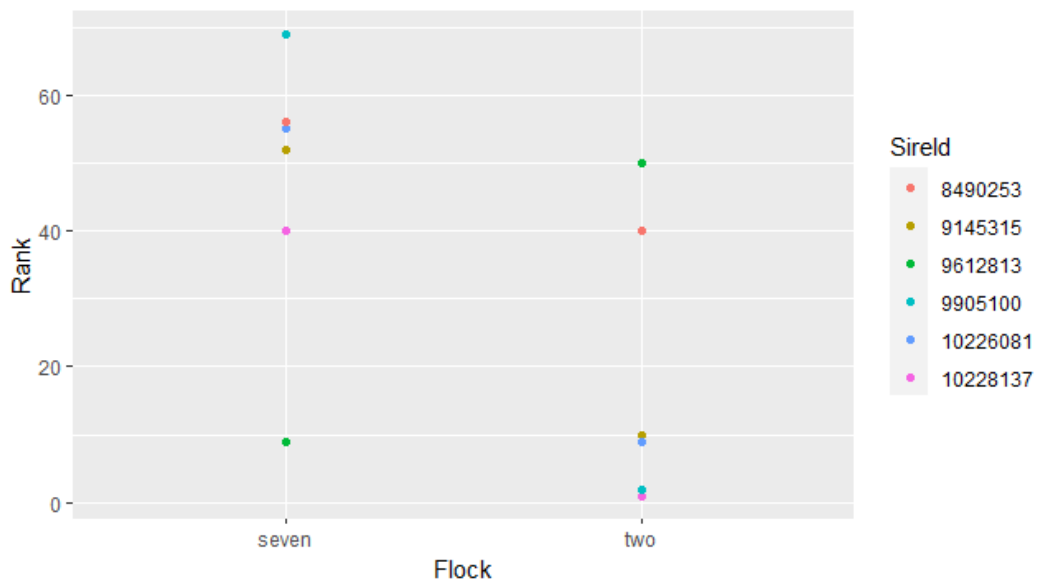
**Figure 4.4** Ranked EBVs for Muscle Depth for the same sires between two flocks showing minimal or no reranking



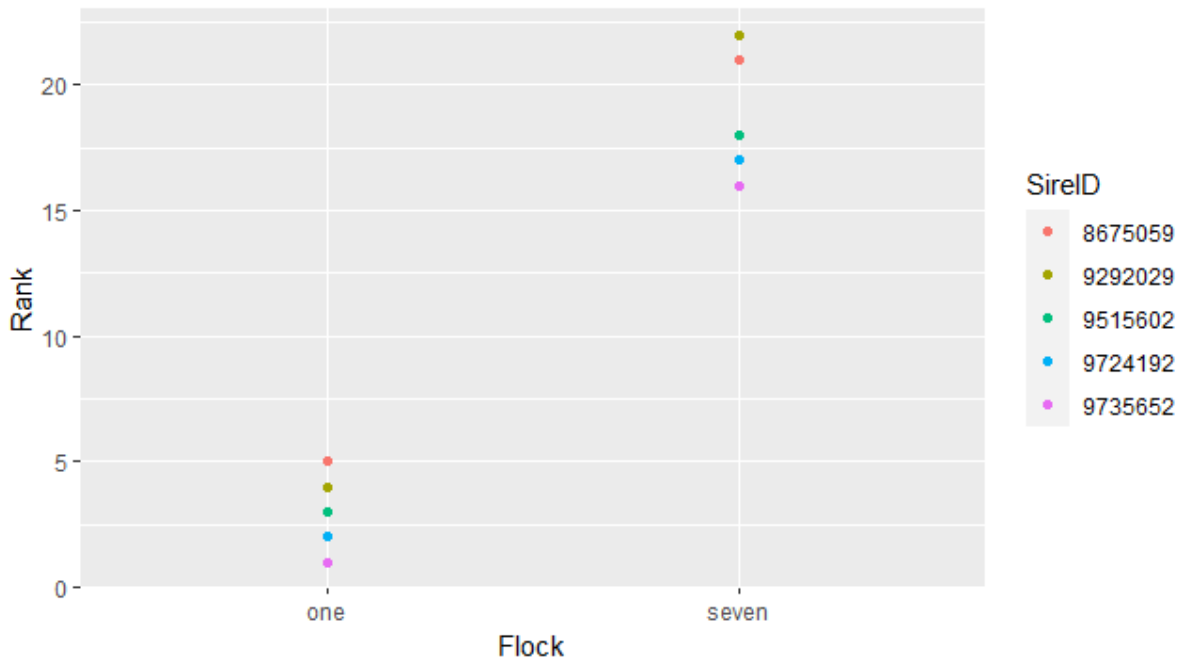
**Figure 4.5** Ranked EBVs for Scan Weight for the same sires between two flocks showing reranking.



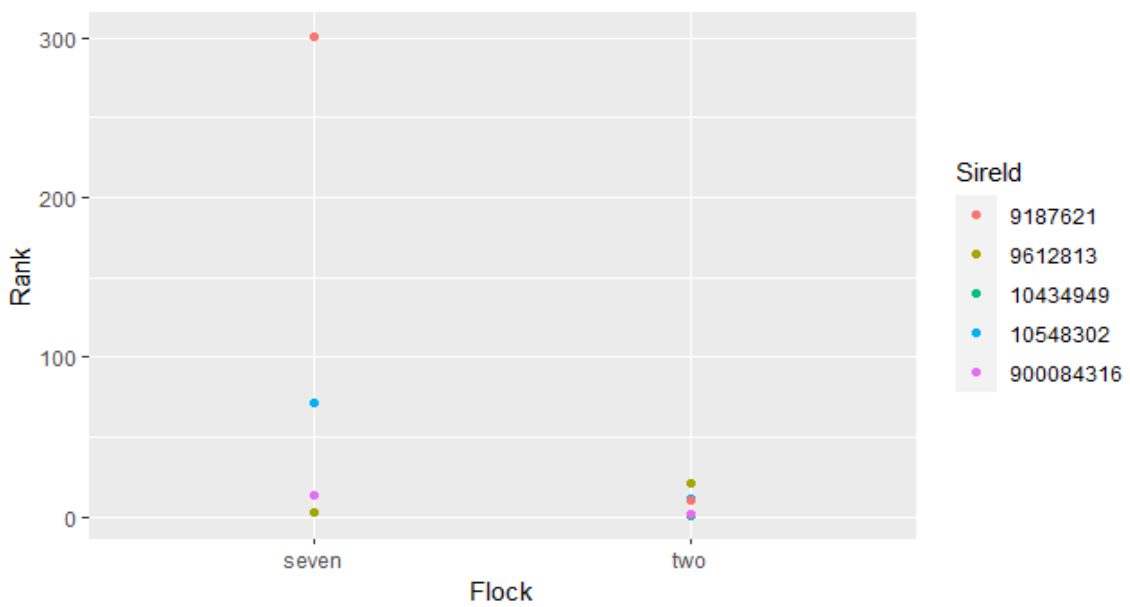
**Figure 4.6** Ranked EBVs for Scan Weight for the same sires between two flocks showing minimal or no reranking.



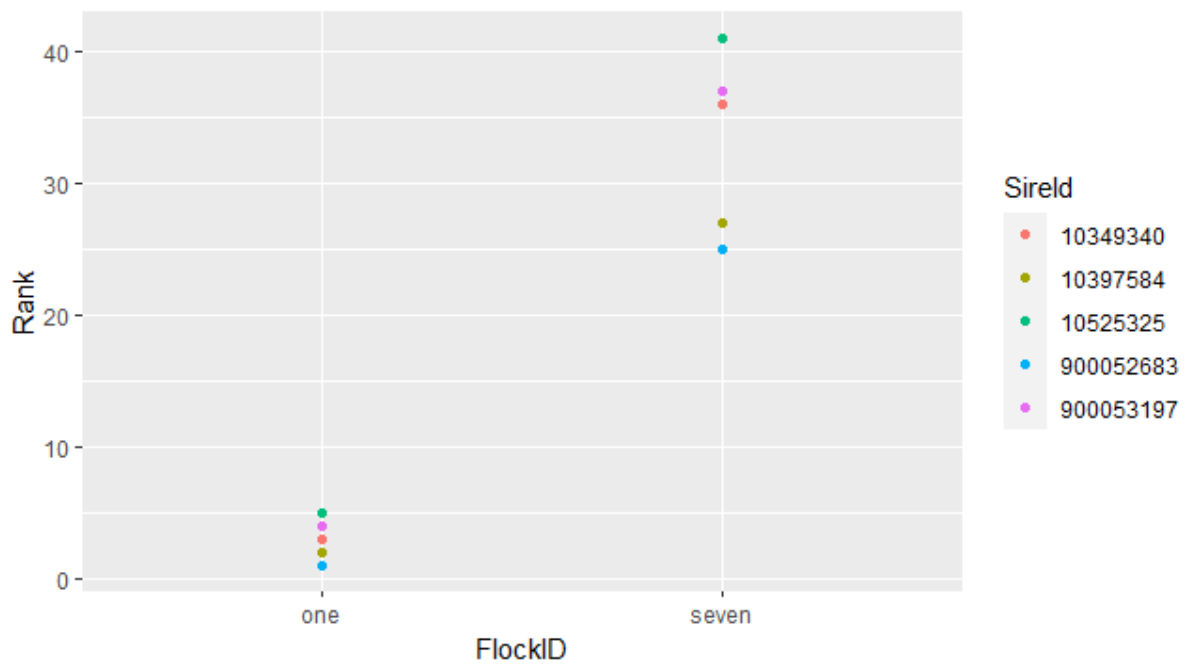
**Figure 4.7** Ranked EBVs for Weaning Weight for the same sires between two flocks showing reranking.



**Figure 4.8** Ranked EBVs for Weaning Weight for the same sires between two flocks showing no or minimal reranking



**Figure 4.9** Ranked EBVs for Birth Weight for the same sires between two flocks showing reranking.



**Figure 4.10** Ranked EBVs for Birth Weight for the same sires between two flocks showing no or minimal reranking.

#### 4.3.2 Reaction Norm Modelling

For GxE analysis using Model 4.2, models for birth weight did not converge while weaning weight, scan weight and muscle depth models did converge. Thus, results could not be displayed for the birth weight model and fat depth; noting that the models for fat depth had previously failed due to lack of records. The GxE effect accounted for 13%, 25% and 26% of the total variance for weaning weight, scan weight, and muscle depth respectively. This is a larger percentage of the total variance than presented in other studies (Maniatis and Pollott, 2003) However, this is likely because the models did not include additional random effects like maternal effect or the covariance between maternal and animal effect which did account for 20% of the phenotypic variance in Maniatis and Pollott (2003). The reaction norm produced

EBVs across the flock deviation for all sires in the pedigree meaning that EBVs represent the breeding value across environments with the GxE effects taken into consideration.

Heritability values were calculated across the available flocks for the three traits. The heritability values are displayed in the Table 4.1 for weaning weight, 4.2 for scan weight and 4.3 for muscle depth.

**Table 4.1** Heritability values across flocks for weaning weight

Flock Number	51459	51461	51462	51463	54161	54163	54572	54576	54915	55774
Heritability	0.34	0.48	0.21	0.34	0.35	0.26	0.19	0.17	0.27	0.27

**Table 4.2** Heritability values across flocks for scan weight

Flock Number	51459	51461	51462	51463	54161	54163	54572	54576	54915	55774
Heritability	0.28	0.17	0.19	0.22	0.28	0.17	0.17	0.18	0.17	0.46

**Table 4.3** Heritability values across flocks for muscle depth

Flock Number	51459	51461	51462	51463	54161	54163	54572	54576	54915	55774
Heritability	0.18	0.16	0.17	0.16	0.17	0.16	0.16	0.17	0.17	0.21

For each trait the top three sires with progeny in the flocks that were analyzed and their EBV are as follows:

**Table 4.4** Top three ranked Sires and EBVs from Reaction Norm Analysis of Muscle Depth

SireID	EBV
8777995	4.425
9625176	3.983
9097015	2.275

**Table 4.5** Top three ranked Sires and EBVs from Reaction Norm Analysis of Weaning Weight

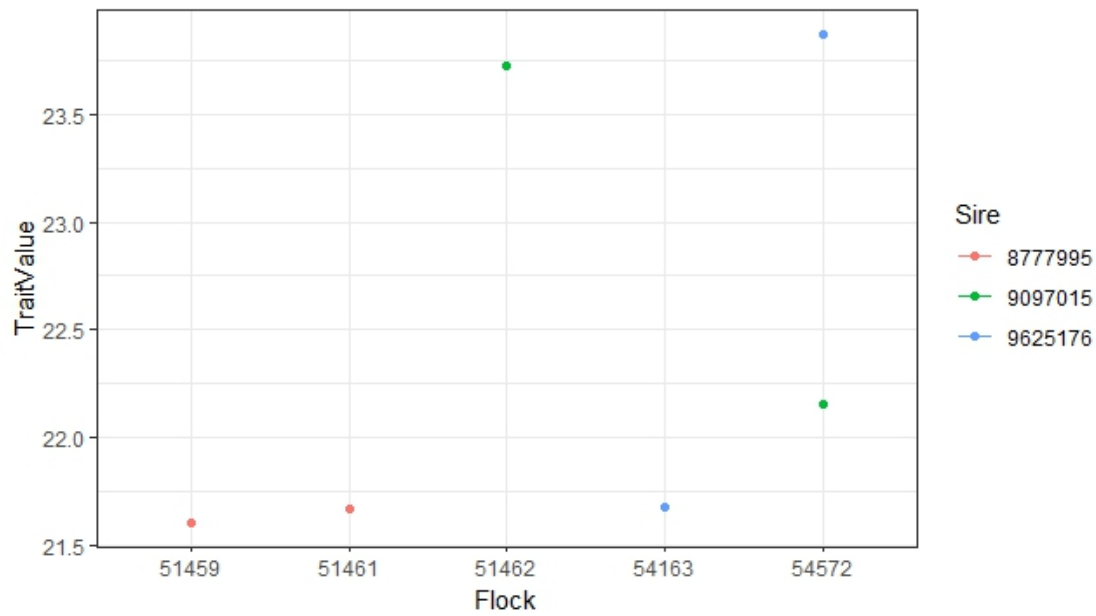
SireID	EBV
9625176	4.591
9193746	3.56
9612813	3.496

**Table 4.6** Top three ranked Sires and EBVs from Reaction Norm Analysis of Scan Weight

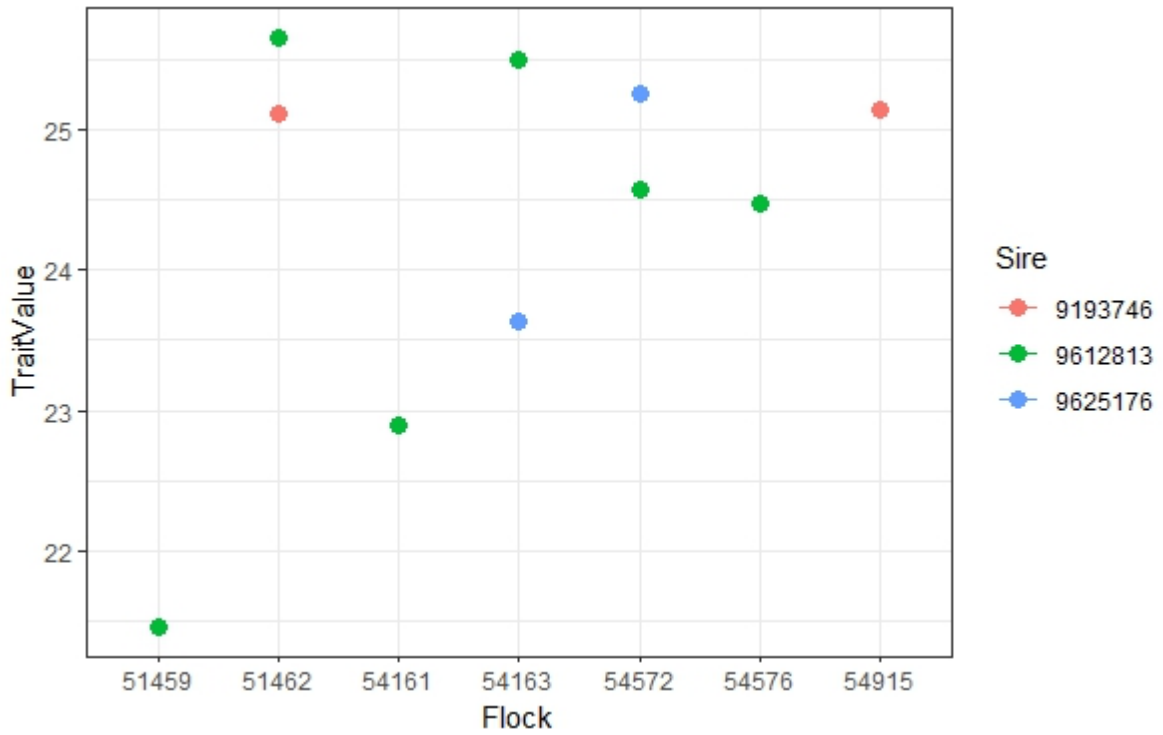
SireID	EBV
8777995	5.879
9625176	5.277
1086209	4.811

Additionally, the top three sires' progeny's phenotypic performance was plotted across the flocks in which their progeny were present. These plots used the phenotypic average of the trait within each flock for each sire to compare how the sires perform across environments. Even though the top sires are the most resilient across environments there are slight changes in the

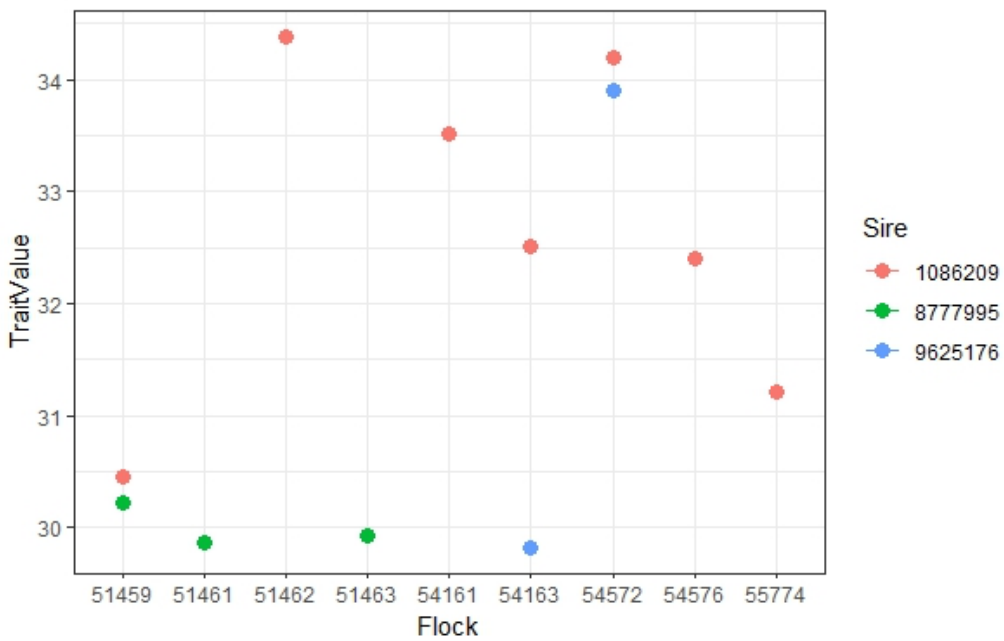
performance of their progeny in different environments which is consistent with the presence of a GxE effect. The plots were constructed in R studio (Rstudio Team, 2012) using “ggplot”.



**Figure 4.11** The phenotypic average for Muscle Depth for the top three sires across flocks.



**Figure 4.12** The phenotypic average for Weaning Weight for the top three sires across flocks.



**Figure 4.13** The phenotypic average for Scan Weight for the top three sires across flocks.

#### 4.4 DISCUSSION

This study into the presence of GxE effects in the terminal sire progeny testing project, RamCompare revealed that although environments across the UK are more similar than comparisons between international environments, there are GxE effects present for the two weight traits, weaning weight and scan weight, and ultrasound muscle depth. These findings are consistent with other studies in GxE effects in UK sheep (Bishop et al., 1996; Maniatis and Pollott, 2003; Bishop et al., 2010; McLaren et al., 2012a; McLaren et al., 2012b). McLaren et al. (2014) found that GxE effects were present in Texel sheep in the UK for weight measured at 21 weeks of age and ultrasound muscle depth. In that study, sire reranking was present suggesting that GxE effects could impact the accuracy of selection for Texel sires across the UK. Additionally, McLaren et al. (2015) found the GxE effects were present for 21 week weight, ultrasound muscle depth and ultrasound fat depth in Charollais sheep in the UK. There is growing evidence that GxE effects exist across UK environments and can impact the accuracy of selection in terminal sire breeding evaluations. This study adds to this growing body of evidence by looking at a cross breed dataset that includes multiple different breeds of sires across the UK. In the study, the GxE effects that were present cause reranking of sires to the point that a sire can be the most well adapted for one flock and the least well adapted for another flock. These effects are consistently shown across the two modelling strategies that were used in the study. However, it should be noted that this could be due to the breed of ewes being used in each flock as well as other maternal effects. The lack of information on ewe breed prevented this study from investigating the impact of ewe breed.

However, the design of the breeding program and the amount of additional environmental data collected limited the conclusions that can be drawn from this study. In terms of the breeding program, given that each sire was only used in two flocks, it was only possible to compare

performance and assess GxE effects in two environments. There may have been more substantial reranking of sires if the same set of sires had been used in every flock. Thus, additional research should be done to assess GxE effects with a breeding program that exposes the sire genetics to more environments. One set of sires should be identified, and those exact same sires should be used across multiple flocks. The results would be able to compare sire performance over all the flocks in the study rather than the two flocks that could be compared in this study. This would offer a more complete understanding of the impact of GxE effects in the UK's commercial sheep industry. Additionally, it would offer RamCompare breeders more information on which sire is best suited to the environmental conditions of their flock.

This study was also limited by the extent of data collected from the RamCompare breeders. Other studies looking at GxE effects in the UK have incorporated additional environmental information including different planes of nutrition (Osoro et al., 1999), grazing management (McLaren et al., 2014), vitamin and mineral administration (McLaren et al., 2014), type of concentrate feed used (McLaren et al., 2015) and management at weaning (Espasandin et al., 2013). In this study the only environmental information collected was climatic information based on the Post Code of the farm, the flock of birth, and the season and year of birth. If additional information including ewe breed, farm type, altitude, latitude or weather data were available this could have been used to create an environmental scale that could be used as the environmental descriptor in the reaction norm models. This would have allowed for greater clarity on how environments differed. Due to a lack of differentiations between environments based on climatic information, the flock and the phenotypic difference was used as the best indicator of environmental variation. The lack of information on other environmental factors prevented this study from investigating the impact of a specific environmental feature on the GxE effects. This may limit the power of this GxE study because flocks could only be

differentiated based on phenotypic performance thus identifying what environmental factors impact performance and to what degree is not possible with these models. However, this study was able to showcase the modelling technique shown in Fikse et al. (2003) using phenotypic deviation across flocks as the environmental descriptor as a useful method especially when there is a lack of additional information on the environmental and management conditions present in each flock.

#### ***4.4.1 Sire and Flock Interaction***

Sire models like that described by Model 4.1 have been used in other studies to assess genotype x environment interactions in sheep primarily using contemporary groups such as flock (Brown et al., 2009; Reddy et al., 2018; Kelman et al., 2022). Including a flock or herd effect has been seen to improve genetic models for weight traits in beef cattle (Robinson, 1996; Lee and Pollak, 1997; Meyer, 1997). However, in order to accurately estimate any GxE effect for a weight trait, each sire must be used in multiple flocks. The limitations of the breeding program in the RamCompare project have already been discussed in this chapter. However, it was still possible to compare sire performance across at least two flocks for the traits of interest.

In the sire model the significant sire effect suggested that the sire's genetic potential plays an important role in the progeny's performance for the traits of interest. This is consistent with what was found by Brown et al. (2009) when they used similar sire models with a flock interaction to assess sire performance in Lambplan. Lambplan is a terminal sire breeding assessment used in Australia that is comparable to RamCompare in that it collects progeny data for the most popular terminal sire breeds in Australia's commercial sheep industry. In this study, the sire and flock effects explained between 2 and 4% of the phenotypic variance and

improved the fit of the model (Brown et al., 2009). The study did not assess the amount of reranking present between sires, but it did suggest that the genetic evaluations benefited from the inclusion of the sire by flock effect.

In this study, the sire models with the contemporary group environmental indicator of flock show levels of reranking which are consistent with the presence of GxE effects (Falconer and Mackay, 1996). Additionally, the sire by flock effect accounted for a significant amount of variance. Reranking was present for all the traits of interest. There were differences in the degree to which reranking was present between the traits. This is likely due to the breeding program and resulting data structure wherein only two flocks could be compared for most of the sires in the dataset. Nonetheless GxE did result in reranked sires for at least some of the flocks in the analysis. Further research should be conducted which could assess the level of reranking when the sires have progeny in more flocks, thereby exposing them to a greater variety of environments. This modelling method is helpful when there are few environmental indicators other than flock as it uses the contemporary group as the representation of environmental variation. Additionally, the results offer an environment specific recommendation to the producer. The sires are ranked within each flock showing which sire is best suited to perform well for the trait of interest within each environment. However, the results from Model 4.1 do not indicate which sire is the best overall or the most resilient across environments. The second set of models described by Model 4.2 offer are more properly suited to assess overall performance across environments.

#### ***4.4.2 Reaction Norm Modelling***

The final set of reaction norm models described by Model 4.2 indicated that GxE effects were present in the terminal sires included in the dataset. The results from this analysis show which sires have the highest genetic potential overall with GxE effect included in the analysis. This means that the sires displayed with the highest EBVs for weaning weight, scan weight and muscle depth should perform well across environments because they are the most resilient to the GxE effects. Additionally, the phenotypic performance of the top three sire's progeny was plotted across flocks showing that there is both scaling and reranking present between sires wherein their progeny's performance changes between flocks. Also, the heritability values for all three traits differ across flocks. The weight traits have greater variability in heritability values compared to the values for muscle depth. This suggests that GxE effects are present for all three traits but have greater impact on the weight traits over muscle depth.

GxE analysis based on average phenotypic performance for a flock or herd has been used successfully in both sheep and cattle assessments (Kolmodin et al., 2002; Fikse et al., 2003; Pollott and Greef, 2004; Strandberg et al., 2009; Santana et al., 2013). It is a particularly good option when there are not many variables that can differentiate environments or when environments are highly similar. In this case the flocks had similar environmental conditions indicated by rainfall and temperature and there were no other environmental indicators except for the contemporary group indicated by the flock variable. Thus, deviations in phenotypic performance for each flock were a more reliable indicator of environmental difference than an environmental cluster.

The issues with model convergence for the reaction norm model for birth weight were not unexpected considering the findings from the previous genetic analysis in Chapter Three where the lacunae of records resulted in high correlations between the intercept and slope of the polynomial in the RC dataset. Additionally, the reaction norm models depended on phenotypic deviation for the trait of interest. For birth weight, the phenotypic deviation between flocks was smaller than for the other traits. The birth weight deviations from the mean ranged 1.036 kg (-0.5780 kg – 0.4580 kg), the weaning weight deviations from the mean ranged over 6.124 kg (-3.180 kg -2.940 kg), the scan weight deviations from the mean ranged over 6.99 kg units (-4.460 – 2.530) and muscle depth ranged over 4.93 kg (-2.400kg – 2kg ). Thus, it may have been difficult for the model to converge due to a lack of variance in the phenotypic deviation. The breeding program structure could have also contributed to the non-convergence because most sires were only used in two flocks which could have reduced the overall variance in sire performance.

However, the models that did converge for scan weight, weaning weight and muscle depth indicated that the reaction norm model with a regression on phenotypic deviation as an environmental descriptor result in a significant GxE effect. This is consistent with other studies that have used a reaction norm model to assess GxE in UK sheep (Bishop et al., 1996; McLaren et al., 2012a; McLaren et al., 2012b; McLaren et al., 2014). McLaren et al. (2015) found evident GxE for 21 week weight and ultrasound muscle depth in their study using both a phenotypic deviation as an environmental indicator and an environmental scale determined primarily by the use of supplementary feed, weather conditions, and weaning management among other factors. The reaction norm modelling portion of this study could provide further insight if more environmental information was incorporated into the dataset allowing for a reaction norm using

an environmental scale in addition to the method using the phenotypic deviation as the environmental indicator.

Additionally, this study also showed that random regression modelling techniques are highly sensitive to the distribution and availability of records across the age or time trajectory. The relationship between distribution and variability of growth data and random regression modelling should be explored further. Previously, random regression modelling has been shown to be sensitive to the number of records in the genetic analysis portion of this thesis. Thus, a similar study with a larger number of sires, consistent availability of records and greater range of environments could potentially aid reaction norm model convergence therefore providing even more insight into GxE effects in the commercial sheep population in the UK.

#### ***4.4.3 Implications for Ramcompare Breeders***

Overall, the study showed that both Model 4.1 and Model 4.2 indicated GxE effects were present and provided useful recommendations for RamCompare breeders even though the breeding program only allowed for comparison between a small number of flock and there was a lack of environmental data. This study offers a path forward to allow future genetic evaluations to consider GxE effects in this widely used commercial sire evaluation. The reduced requirement for environmental information in this study which used a phenotypic deviation rather than a scale based on clustered or grouped environmental information may be of use to similar commercial sheep assessments where environmental information is difficult to collect or not widely reported. Although an environmental scale may be able to compare the effect of more specific environmental conditions and management practices on a particular trait

of interest, the commercial sheep industry can still account for GxE effects in terminal sire assessments without taking on a much higher level of data collection.

Additionally, the two methods used in this study provided specific results that could assist producers in two separate ways. Model 4.1 showed which sires performed best for each trait within each flock. If a breeder or producer was looking to specifically optimize his or her flock's performance for muscle depth or scan weight, then these models offered him or her a specialized recommendation for the flock and trait of interest. Therefore, the study suggests that the RamCompare project should analyze a sire's performance within the environment the producer wants to use. Thus, taking the GxE effect into account is necessary to produce reliable genetic information and recommendations on the appropriate sire for a given breeder or producer.

Secondly, Model 4.2 provides an additional set of recommendations to breeders and producers. If they are looking not for the sire that will perform the best in one flock or another but the best overall, to incorporate into a national breeding program for example, then the reaction norm model can provide those recommendations for each trait. Thus, if a breeder or producers lacks information on the flock of interest or wants to use a sire in more than one flock, the reaction norm model can account for the GxE effect and recommend a sire for each trait that has the most resilient performance record across flocks. Thus, in both cases, the RamCompare project should include GxE effect into its analysis in order to produce the most accurate EBVs. The two methods used in this study do not require additional data on the environment and therefore require little additional effort from RamCompare breeders while still offering a genetic assessment that accounts for the GxE effects present in the RamCompare project's commercial sire evaluation.

## CHAPTER FIVE: GENERAL DISCUSSION

The overall aim of this thesis was to investigate the applicability of regression modelling techniques for growth modelling and subsequent genetic evaluations in commercial sheep in the UK. Individual studies examined growth modelling, relating growth to non-longitudinal traits, optimizing growth and carcass trait selection, and finally assessing GxE effects. This modelling technique, using the infinite dimensional methodology, solves some problems in multi-trait methodologies including impracticality, neglect of random variability in environmental effects and the risk of overparameterization. Random regression modelling has been used with success for evaluations of longitudinal data in other species especially in dairy cattle evaluations. However, the models were largely developed on lactation evaluations where there are multiple records repeated across several years that produce repeated curves over the animal's lifetime. In comparison, sheep growth curves generally have fewer records taken throughout the animal's life because sheep growth produces only one curve from birth to slaughter. Thus, there are some key differences in using random regression models for growth evaluations. The key determinants in the success of the growth models are the number of times an animal's weight is recorded, when those recordings are taken and how the records should be distributed across the age trajectory.

With these considerations in mind, random regression models have been used in genetic evaluations for growth in sheep (Fosscoco and Notter, 1995; Lambe et al., 2006; Sallam et al., 2019). Many of these successful genetic evaluations are used in highly managed flocks where random variability may be better accounted for than in commercial environments. There are far fewer studies that investigate random regression models in commercial environments. This thesis sought to fill this gap by applying random regression methods to commercial sheep

production. After investigating the applicability of random regression modelling for genetic evaluations in commercial sheep breeds, the results in this thesis suggest that random regression models are highly sensitive to the distribution of weight records across the age distribution. When there are few records across the distribution, or these records are not evenly distributed, this can lead to high correlations between the model parameters and issues with model convergence (McGowan et al., 2022). In commercial flocks, weight records are rarely equally distributed across the age distribution. Thus, when random regression models are used in commercial sheep evaluations, the models must find ways to account for the difficulties in adequately modelling the growth curve when records are sparse and obtain appropriate genetic parameters. In the following sections of this chapter each aim of the thesis will be discussed in turn and conclusions will be drawn across the three studies included in the thesis.

## **5.1 USING RANDOM REGRESSION MODELLING FOR GROWTH IN SHEEP**

In the second chapter of the thesis, random regression models were used to model growth in Suffolk and Charollais sheep. In both datasets there were four possible weight records that could have been taken with a significant lack of records at the mature end of the age distribution. In initial modelling attempts, the genetic parameters for the weight trait were inflated beyond biological plausibility. It was observed that in the random regression model, the parameters of the model were highly correlated with a value of 0.99. When this correlation was fixed, models for both sheep breeds produced biological parameters for growth that were consistent with published literature estimates. The same phenomenon occurred in Chapter Three of the thesis when Scottish Blackface and RamCompare datasets were used for growth modelling.

These datasets had the same four possible weight records with Scottish Blackface data having records available at all points. In the RamCompare data, on the other hand, there are fewer weight records as the age increases. Again, fixing the correlation between the slope and intercept of the Legendre polynomial at a level of 0.5 produced reliable genetic parameters that corresponded with published literature estimates for heritability of growth between 0.1 – 0.4 (Fosecco and Notter, 1995; Lambe et al., 2006; Amer et al., 2007; Arthy et al., 2020) . These results suggest that a distribution with four or fewer weight records that are not evenly distributed across the age distribution causes inflated correlations between the parameters of the Legendre Polynomial used in the random regression. These inflated correlations result in improbable genetic parameters and inaccurate EBVs. Thus, to successfully model and evaluate growth in commercial sheep production, more weight records should be taken across the animal's life. The Scottish Blackface dataset was collected from a research flock rather than a commercial flock. This was done to allow a preliminary comparison between commercial flocks and research flocks. If the commercial environment alone was creating wide random variability or the sparsity of records for mature weight was solely responsible for the inflated correlations in the random regressions than one would have expected the research flock to perform differently than the commercial flocks. In Chapter Three, the Scottish Blackface dataset had an inflated correlation that was fixed to produce reliable genetic parameters. Thus, it may not be the commercial environment but the total number of weight records or the unequal distribution of the records across the animals' lives that cause inflated correlation values. Further studies should be conducted using commercial sheep data and research data with more weight records across the distribution or using simulated data with weight records across the age distribution to further investigate the impact the commercial environment may have on genetic evaluations.

Although fixing the correlation between the parameters of the Legendre Polynomial resulted in more accurate genetic parameters for all the growth models in this thesis, it could have introduced additional bias. Thus, this thesis includes a novel validation methodology to determine if the fixed correlation of 0.5 predicted accurate EBVs. In the validation procedure, each set of EBVs produced by the models with fixed correlations at low, medium and high levels were compared to the corrected yields which represent the animals observed genetic merit; these corrected yields were calculated as the genetic merit of each individual animal for the growth trait after it had been corrected for environmental variance in the form of fixed effects and residual variance. The approach, although pragmatic, did suggest that constraining the correlation between the parameters of the random regression model could allow for random regression modelling and biologically reasonable genetic evaluation of growth in commercial sheep flocks where data distribution is irregular, or data availability was reduced.

Further research investigating the appropriate distribution and number of weight records for random regression modelling of growth should be conducted to identify the appropriate number of weight records for the models. Random regression modelling has the potential to provide more accurate genetic parameters with more flexible data collection procedures (Lambe et al., 2004). However, the number of records in these datasets was not sufficient for random regression models without fixed parameter correlation values. However, the fixed correlation method does offer a novel methodology to employ in other commercial growth models where data may be sparse. The fixed correlation models did produce biologically plausible results without introducing so much bias into the model that the genetic predictions were untenable. In a sense, the approach could be regarded as bending the genetic covariance matrix for the regression coefficients in the estimation process to ensure covariances are positive definite.

Thus, the first aim of this thesis which was to assess the possibility of random regression modelling for growth for commercial sheep population in the UK was achieved. The models with fixed correlations produced EBVs that were consistent with observed genetic merit. This thesis contributes to the growing body of research on the use of random regression methodologies in commercial settings where data distribution may be imperfect by suggesting remedies for problems that may arise like the inflated correlation values. Additionally, achieving this aim provides practical recommendations for the commercial flocks in the study.

The heritability and genetic correlations of the growth trait suggest that it is moderately heritable across the entire age distribution and can therefore be influenced through genetic selection. The results also suggest that the growth trait is consistently genetically influenced but weights that are far apart on the age distribution are less highly correlated meaning that focusing on improving weights at a certain point on the age distribution will have a larger impact on weights close on the distribution than on those further away. Finally, if further evaluations are done on these commercial populations, more weight records should be collected, and the resulting genetic parameters compared to those found in the studies contained in this thesis to compare the impact of data availability on the genetic parameters produced by the models.

## **5.2 RELATIONSHIP BETWEEN GROWTH AND NON-LONGITUDINAL TRAITS**

The second aim of the thesis was to relate body weight to non-longitudinal traits with a special interest into those non-longitudinal traits that are candidates to be used in future sheep

evaluations. The current sheep evaluation focuses on traits that are used in the assessment and scoring of carcasses. In order to achieve this aim, Chapter Three used a dataset with Scottish Blackface sheep that had records for carcass weight, conformation and fat class. Slaughter weight was included in the selection indices for carcass traits but was better categorized as a live weight.

The analysis of relationship between growth and non-longitudinal traits centred on the bivariate random regression analysis of growth and the four carcass traits that were included in the dataset (Lewis et al., 2004; Macfarlane et al., 2004a; Macfarlane et al., 2004b; Massender et al., 2019). Only two of the four models converged which was not particularly surprising considering the problems with the univariate random regression models for body weight (McGowan et al., 2022). Thus, investigating the relationship between growth and non-longitudinal traits was confined to conformation and fat class. There was a strong positive genetic correlation between body weight and both fat class and conformation respectively. This was consistent with other studies investigating the relationship between body weight and carcass trait (Lewis et al., 2004; Macfarlane et al., 2004a). However, the phenotypic relationship between body weight and these two non-longitudinal traits was less clear. There was a small phenotypic correlation between the body weight and conformation while fat class and body weight have a more consistent positive phenotypic correlation that increases as age increases.

The aim to establish a relationship between body weight and non-longitudinal traits was achieved in the results from these models. It serves as a preliminary fact which leads directly to the third aim of the thesis which is to optimize body weight selection for the most beneficial impact on the carcass trait.

### **5.3 OPTIMIZING SELECTION CRITERIA FOR GROWTH AND CARCASS TRAITS**

The optimization of body weight selection for impact on carcass traits has direct practical impact for commercial sheep producers. Carcass traits are the largest determinants of the profit sheep producers make when their animals are sold for processing (Lambe et al., 2007, 2008). If selection for or focus on an animal's weight at a specific point on the growth curve can optimize the carcass trait this would be a direct improvement for the profitability of commercial sheep production.

A preliminary selection index was constructed to assess how accurately a carcass trait can be predicted from the body weights and which weight has the largest impact on the carcass trait. The selection indices generally showed that slaughter and carcass weight traits could be predicted from body weights recorded during the animals' lives with reasonable accuracy, which was consistent with other selection indices using body weight (Simm and Dingwall, 1989; Marquez et al., 2013). Specifically, Marquez et al., (2013) found that the lean sire index was successful in producing heavier lambs or leaner lambs at finishing, to fit the aims of the breeder, thereby influencing the carcass weight across three major terminal breeds.

In this study, for both carcass traits the fourth weight had the most impact on the carcass trait in the selection index suggesting that if producers want to increase this their slaughter or carcass weight then they should increase the weight recorded immediately prior to slaughter. Clearly,

decreasing this weight would also decrease the slaughter and carcass weight. However, it should be noted that slaughter weights and carcass weights are usually prescribed to the producer, thus achieving a specific target in a short amount of time is more important than a simple increase or decrease in one of the live weights recorded during the growth cycle. Thus, it may be more helpful to focus more on a steep growth curve representing a fast growth rate rather than a single weight being the most impactful on the slaughter or carcass weight.

However, the non-longitudinal traits which had been shown to have a genetic relationship with body weight, in the bivariate random regression models, were much less accurately predicted from body weight in the selection index. The less accurate selection indices for conformation and fat class also provided conflicting results on which weight had the largest impact on the carcass trait. The conflicting results could be a consequence of the data structure or because these two variables are categorical variables rather than continuous numeric values. Although the selection index did not provide clear optimization for conformation or fat class, the bivariate random regression models did have useful results. The consistent positive genetic relationship between the carcass traits and body weight suggests that focusing on improving body weight closest to slaughter will have the largest effect on the carcass trait. This is consistent with Lambe et al., (2008) establishing that CT measured fat was genetically correlated to back fat depth at weaning and the fat score received at slaughter. This suggest that tissue wights, which are one of the components of body weights, do have an impact on the fat score received at slaughter. Similarly selecting on the basis of conformation score has a relationship with muscularity measured at the hind leg and lumbar regions of the animals (Lambe et al., 2008).

However, these models did present inflated heritability values for the weight trait. The heritability values in these models are higher than the previous models with body weight only and they are outwith the limits of published literature estimates for the growth trait. Again, the number of weight records available could impact the bivariate random regression models. Additionally, these models did not account for tissue weights rather than total body weight as Lambe et al., (2008) included nor did they account specifically for the seasonal fluctuations in tissue levels which can impact growth as in Lambe et al., (2005) nor could they include the impact of multiple production cycles as in Lambe et al., (2004) as this information was not included in the datasets and the models struggled to converge with the available variance components.

Nonetheless, the aim of optimizing selection on body weight for positive impact on carcass trait was achieved through two modelling techniques. Additionally, practical recommendations can be provided to UK sheep producers; namely to focus on optimizing weights close to slaughter to optimize the carcass trait. However, this study was limited by the data structure of the Scottish Blackface dataset wherein the distribution of records across the age range was minimal. The study was also limited by the management of the Scottish Blackface flocks wherein animals are identified for slaughter once they achieve a target live weight and then do not have weight recordings for the remaining weight recording periods. Again, similar studies should be conducted with more weight records that are distributed across a larger range of the age distribution. This could provide more reliable results with heritability values that are more consistent with published literature estimates.

## 5.4 ESTIMATE GXE INTERACTIONS IN UK SHEEP BREEDS

The final aim of the thesis was to investigate the presence of GxE effects in commercial UK sheep populations. The project used data from RamCompare, the terminal sire progeny testing project conducted by AHDB. Using these data means the recommendations from this study have wide significance for the UK sheep industry. The two methodologies used in the study contributed to the assessment of GxE effects and provided practical recommendations for sheep breeders. The models that used the flock where the sire's progeny was born as an environmental indicator showed a significant interaction between the terminal sires and the flock. This was consistent with other studies showing that GxE effects are present when sire and contemporary group effects are included in models for growth (Lee and Pollak, 1997; Brown et al., 2009). When the EBVs were calculated for the sires within each flock and then ranked there was significant reranking of sires. There was scaling of EBVs across flocks for many of the sires. This is expected where GxE effects are impacting the sire's performance. Practically speaking this means that the RamCompare project includes terminal sires that perform well in some environments and poorly in others even though the environments across the UK are not easily differentiated by climatic characteristics.

The second set of models for this study used random regression techniques through a reaction norm model with a regression on each flock's phenotypic deviation from the trait mean. The phenotypic deviation is an excellent option for the reaction norm model when environments are not easily differentiated from one another. In this case, the environments in the UK were largely similar based on the average temperature and rainfall of each flock. Thus, the phenotypic deviation served as a simple way to describe environmental differences. Models

including phenotypic deviation successfully modelled GxE effects in cattle and sheep (Kolmodin et al. 2002, Fikse et al., 2003 Pollott and Greef 2004, Strandberg et al., 2009 Santana et al., 2013). Unfortunately, the reaction norm model for birth weight did not converge. This could be a result of the data structure where the phenotypic deviations were not available for all animals in the dataset but given the issues the dataset presented in previous regression modelling it was not altogether surprising.

The models that did converge for weaning weight, scan weight, and muscled depth showed that the GxE effect accounted for a significant amount of the variance in the model. The resulting EBVs and top three sires are presented as the highest performing sires across all the environments meaning that these sires are the most resilient to the GxE effects. They are the most likely to perform well in a myriad of environments. However, it should be noted that the top sires for each trait are different suggesting that selecting one of these traits may not optimize all of them (Bishop et al., 1996; McLaren et al., 2012a; McLaren et al., 2012b; McLaren et al., 2014.) Further work could be done to understand how well the top ranking sires for one trait rank for the others in order to enhance the recommendation for sheep breeders and producers. The sires that perform the best overall may not be the best performing sire for a specific flock. Thus, the producers' goals determine the recommendation this study provides. If a producer wants the highest performing sire for their environment, even though that same sire's progeny may perform poorly elsewhere, then the model using the sire and flock interaction can indicate which sires are best for their flock. However, if a producer is interested in resilient progeny that will perform well, though potentially not the best, across several environments then the reaction norm models provide the top three sire recommendations for each trait of interest. It should be noted that the reaction norm models were limited by the previously established

relationship between data availability, distribution, and model convergence. This relationship should be explored further wherein a similar study is conducted using a larger number of sires, consistent availability of records and environments that are more clearly differentiated by climatic characteristics. The final aim of this thesis was achieved in the study and the GxE effect present in the RamCompare project may significantly impact the recommendations the project provides for the breeders in the program.

Overall this thesis achieved its aims though the methodologies were highly impacted by the limitations of commercial data structure. Following from this research a few key areas are identified for further study. The first area is the adequate number and necessary distribution of weight records for random regression modelling. Secondly, the potential impact of commercial environments compared to research environments on genetic evaluations should be investigated further. Finally, GxE effects should be further assessed in the RamCompare project where progeny have consistent weight records and sires have progeny in a larger number of flocks especially for birth weight. This thesis had provided greater insight into the impact of data structure on random regression modelling, the importance of tailoring models for commercial data environments, the relationship between longitudinal body weight traits and non-longitudinal carcass traits and the presence of GxE effects in prominent sires in the UK commercial sheep industry.

## BIBLIOGRAPHY

- Agriculture and Horticulture Development Board 2021. UK sheep facts and figures. Retrieved on 5 September 2022, from <https://ahdb.org.uk/uk-sheep-facts-and-figures>.
- Agriculture and Horticulture Development Board 2017 The characteristics of high performing beef and sheep farms in Great Britain. Retrieved 12 Dec 2020 from: <https://ahdb.org.uk/knowledge-library/the-characteristics-of-high-performing-farms-in-the-uk>.
- Albuquerque, LG and Meyer, K 2001 Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. *J. Anim. Sci.* 79, 2776–2789.
- Amer, R, Emmans, GC and Simm, G 1997. Economic values for carcass traits in UK commercial beef cattle. *Livestock Production Science* 51, 267–281.
- Amer, P.R., Nieuwhof, G.J., Pollott, G.E., Roughsedge, T., Conington, J. and Simm, G. 2007 Industry benefits from recent genetic progress in sheep and beef populations. *Animal*, 1, 1414–1426.
- Arthy, V, Venkataramanan, R, Sivaselvam, SN, Sreekumar, C and Balasubramanyam, D 2020. Suitability of random regression models for growth of Madras Red sheep under a field performance recording system. *Small Ruminant Research* 193, 106260.
- Ashfield, A, Wallace, M, Prendiville, R and Crosson, P 2014. Bioeconomic modelling of male Holstein-Friesian dairy calf-to-beef production systems on Irish farms. *Irish Journal of Agricultural and Food Research* 53, 133–147.
- Bishop SC, Conington J, Waterhouse, A and Simm G 1996. Genotype x environment interactions for early growth and ultrasonic measurements in hill sheep. *Animal Science* 62, 271-277.
- Boettcher PJ, Fatehi J and Schutz MM 2003. Genotype x environment interactions in conventional versus pasture-based dairies in Canada. *Journal of Dairy Science* 86, 383-389.
- Bonny SP, Pethick DW, Legrand I, Wierzbicki J, Allen P, Farmer LJ, Polkinghorne RJ, Hocquette JF and Gardner GE 2016. European conformation and fat scores have no relationship with eating quality. *Animal* 10(6), 996-1006.
- Brickell JS, McGowan MM, Wathes DC 2009. Effect of management factors and blood metabolites during the rearing period on growth in dairy heifers on UK farms. *Domest Anim Endocrinol.* 36(2):67-81.
- Brown, DJ, Swan, AA, Johnston, DJ and Graser, HU 2009. Sire by flock year interactions for body weight in poll dorset sheep. *Proc Assoc Advmt anim Breed. Genet* 18:48-51.
- Bullock, KD, Bertrand, JK and Benyshek, LL 1993. Genetic and environmental parameters for mature weight and other growth measures in Polled Hereford cattle. *Journal of animal science* 71, 1737–1741.

- Coffey, MP, Hickey, J and Brotherstone, S 2006. Genetic aspects of growth of Holstein-Friesian dairy cows from birth to maturity. *Journal of Dairy Science* 89, 322–329.
- Coffey MP, Simm G, Hill WG and Brotherstone S 2003 Selection for energy balance in dairy cows. *Interbull Bulletin* 0: 107.
- Conington, J., Bishop, S.C., Waterhouse, A.W. and Simm, G. 1995 A genetic analysis of early growth and ultrasonic measurements in hill sheep. *Animal Science*, 61, 85-93.
- Elizalde, HF, Carson, A F, Munoz, C 2019. Effects of sire genotype on lamb performance at weaning in extensive sheep systems. *Animal* 13, 213-220.
- Espasandin AC, Urioste JI, Naya H and Alencar MM 2013. Genotype× production environment interaction for weaning weight in Angus populations of Brazil and Uruguay. *Livestock Science* 151, 264–270.
- FAS, 2024 Selecting Finishing Lambs – A practical Guide, Retrieved 1 Feb 2025, <https://www.fas.scot/article/selecting-finished-lambs-a-practical-guide/#:~:text=The%20EUROP%20grid%20considers%20the,the%20higher%20the%20fat%20cover.>
- Falconer DS and MacKay TFC 1996. *Introduction to Quantitative Genetics*, 4<sup>th</sup> ed. Pearson Education Limited, Essex, United Kingdom.
- Fiems LO, Ampe B 2014 Importance of dam BW change and calf birth weight in double-muscling Belgian Blue cattle and its relationship with parity and calving interval. *Animal* 9,94–103.
- Fiems LO 2012 Double muscling in cattle: Genes, husbandry, carcasses and meat. *Animals* 2,472–506.
- Fikse WF, Rekaya R and Weigel KA 2003. Assessment of environmental descriptors for studying genotype by environment interaction. *Livestock Production Science* 82, 223-231.
- Fossceco, S L and Notter, DR 1995. Heritabilities and genetic correlations of body weight, testis growth and ewe lamb reproductive traits in crossbred sheep. *Animal Science* 60, 185–195.
- Gilmour, AR, Gogel, BJ, Cullis, BR and Thompson, R 2016. *ASReml User Guide Release 4.0* VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.

Guo, Z, Lund, MS, Madsen, P, Korsgaard, I and Jensen, J 2002 Genetic parameter estimation for milk yield over multiple parities and various lengths of lactation in Danish Jerseys by random regression models. *J. Dairy Sci.* 85:1596–1606.

Guo Z and Swalve HH 1995 Comparison of different lactation curve sub-models in test day models. *Interbull Bulletin* 0: 75.

Held, J, Slyter, AL, Read, B and Long, B 1997. Effect of Sire Selection on Lamb Growth and Carcass Traits. *South Dakota Sheep Field Day Proceedings and Research Reports, Paper 5, Brookings, South Dakota, USA.*

Hermesch, S, Kanis, E and Eissen, JJ 2003. Economic weights for feed intake in the growing pig derived from a growth model and an economic model. *Journal of Animal Science* 81, 895–903.

Huquet, B, Leclerc, H and Ducrocq, V 2012 Modelling and estimation of genotype by environment interactions for production traits in French dairy cattle. *Genet Sel Evol* 44, 35.

IBIS World, 2024 Sheep Industry Statistics <https://www.ibisworld.com/>.

ILRI 2019. Options for the Livestock Sector in Developing and Emerging Economies to 2030 and Beyond Meat: The Future Series. World Economic Forum Geneva, Switzerland.

Jakobsen, J H , Dürr, J W , Jorjani, H , Forabasco, A, Loberg, A and Philipsson, J 2009. Genotype by environment interactions in international genetic evaluations of dairy bulls. *Proc 18th Assoc Advmt Anim Breed Genet, AAAGB*, 133–142, Roseworthy, Australia.

Jamrozik, J, Schaeffer, LR and Weigel, KA 2002 Genetic evaluation of bulls and cows with single- and multiple-country test-day models. *J. Dairy Sci.* 85, 1617–1622.

Jackson, SP, Miller, MF and Green, RD 1997 Phenotypic characterization of rambouillet sheep expressing the callipyge gene:II. Carcass characteristics and retail yield. *Journal of Animal Science*, 75, 125–132.

Jannoune, A, Boujenane, I, Falaki, M and Derqaoui, L 2015. Genetic analysis of live weight of Sardi sheep using random regression and multi-trait animal models. *Small Ruminant Research* 130, 1–7.

Jones, HE, Amer, PR, Lewis, RM and Emmans, GC 2004. Economic values for changes in carcass lean and fat weights at a fixed age for terminal sire breeds of sheep in the UK. *Livestock Production Science* 89, 1–17.

Kalaldehy Al, M, Swaminathan, M, Gaundare, Y, Joshi, S, Aliloo, H, Strucken, EM, Ducrocq, V and Gibson, JP 2021. Genomic evaluation of milk yield in a smallholder crossbred dairy production system in India. *Genetics Selection Evolution* 53, 73.

Karthik D, Suresh J, Reddy YR, Sharma GRK, Ramana JV, et al. 2021. Farming systems in sheep rearing: Impact on growth and reproductive performance, nutrient digestibility, disease incidence and heat stress indices. *PLOS ONE* 16(1): e0244922. <https://doi.org/10.1371/journal.pone.0244922>

Kelly, HR, Browning, HM, Day, JEL., Martins, A, Pearce, GP, Stopes, C and Edwards, SA 2007. Effect of breed type, housing and feeding system on performance of growing pigs managed under organic conditions. *J. Sci. Food Agric.* 87, 2794–2800.

Kelman, K, Alston-Knox, C, Pethick, DW and Gardner, GE 2022. Sire Breed, Litter Size, and Environment Influence Genetic Potential for Lamb Growth When Using Sire Breeding Values. *animals* 12(4), 1-15.

Kearney JF, Schultz MM, Boettcher PJ and Weigel KA 2004a. Genotype x environment interaction for grazing versus confinement. I. Production traits. *Journal of Dairy Science* 87, 501-509.

Kearney JF, Schultz MM and Boettcher PJ 2004b. Genotype x environment interaction for grazing versus confinement. II. Health and production traits. *Journal of Dairy Science* 87, 510-516.

Kirkpatrick, M and Heckman, N 1989. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. *Journal of Mathematical Biology* 27, 429–450

Kirkpatrick M and Heckman N 1989 *Mathematical Woolly. Evolution*: 429–450.

Kolmodin, R, Strandberg, E, Danell, B and Jorjani, H 2004. Reaction norms for protein yield and days open in Swedish red and white dairy cattle in relation to various environmental variables. *Acta Agriculturae Scandinavica, Section A. Animal Science* 54 (3), 139-151.

Lalit, Malik, ZS, Dalal, DS, Dahiya, SP, Patil, CS and Dahiya, R 2016. Genetic analysis of growth traits in Harnali sheep. *Veterinary World* 9, 128-132.

Lambe NR, Simm G, Young MJ, Conington J, Brotherstone S. 2004. Seasonal changes in tissue weights in Scottish Blackface ewes over multiple production cycles. *Animal Science*. 79(3):373- 385.

Lambe NR, Brotherstone S, Young MJ, Conington J, Simm G. 2005. Genetic relationships between seasonal tissue levels in Scottish Blackface ewes and lamb growth traits. *Animal Science*. 81(1):11-21.

Lambe, NR, Navajas, E A, Simm, G and Bünger, L 2006. A genetic investigation of various growth models to describe growth of lambs of two contrasting breeds. *Journal of animal science* 84, 2642–2654.

Lambe NR, Navajas EA, McLean KA, Simm G and Bünger L 2007. Changes in carcass traits during growth in lambs of two contrasting breeds, measured using computer tomography. *Livestock Science* 107, 37-52

Lambe NR, Conington J, Bishop SC, McLean KA, Bünger L, McLaren A and Simm G 2008. Relationships between lamb carcass quality traits measured by X-ray computed tomography and current UK hill sheep breeding goals. *Animal* 2, 36-43.

Lee, C. and Pollak, EJ 1997 Influence of inbreeding depression on genetic (co)variance and sire-by-year interaction variance estimates for weaning weight direct-maternal genetic evaluation. *Asian-Aus. J. Anim. Sci.* 10:510-513.

Legarra A, Misztal I and Bertrand JK 2004 Constructing covariance functions for random regression models for growth in Gelbvieh beef cattle. *Journal of Animal Science* 82, 1564–1571.

Lewis RM, Macfarlane JM, Simm G and Emmans GC 2004. Effects of food quality on growth and carcass composition in lambs of two breeds and their cross. *Animal Science* 78, 355-367.

Lewis, RM and Brotherstone, S 2002. A genetic evaluation of growth in sheep using random regression techniques. *Animal Science* 74, 63–70.

Macfarlane JM, Lewis RM and Emmans GC 2004a. Effects of two dried forages, and a choice between them, on intake, growth and carcass composition in lambs of two breeds and their cross. *Animal Science* 78, 485-496.

Macfarlane JM, Lewis RM and Emmans GC 2004b. Growth and carcass composition of lambs of two breeds and their cross grazing ryegrass and clover swards. *Animal Science* 79, 387-396.

Maniatis, N. and Pollott, G.E. (2002a) Genotype-by-environment interactions in lamb weight and composition traits. *Animal Science*, 75, 3-14.

Maniatis, N. and Pollott, G.E. (2002b) Nuclear, cytoplasmic and environmental effects on growth, fat and muscle traits in Suffolk lambs from a sire referencing scheme. *Journal of Animal Science*, 80, 57-67.

Maniatis, N. and Pollott, G.E. (2002c) Maternal effects on weight and ultrasonically measured traits of lambs in a small closed Suffolk flock. *Small Ruminant Research*, 45, 235-246.

Maniatis, N. and Pollott, G.E. (2003) The impact of data structure on genetic (co)variance components of early growth in sheep, estimated using an animal model with maternal effects. *Journal of Animal Science*, 81, 101-108.

Marquez, G.C., Haresign, W., Davies, M.H., Roehe, R., Bunger, L., Simm, G. and Lewis, R.M. 2013 Index selection in terminal sires improves lamb performance at finishing. *Journal of Animal Science*, 91, 38-43.

Marshall DJ, Walker RI, Cullis BR, Luff MF. 1991 The effect of footrot on body weight and wool growth of sheep. *Aust Vet J.* Feb;68(2):45-9.

Massender, E, Brito, L, Canovas, A, Baes, C, Kennedy, D and Schenkel, F 2019. A genetic evaluation of growth, ultrasounds and carcass traits at alternative slaughter endpoints in crossbred heavy lambs. *J Anim Sci* 97(2), 521-535.

Maxa, Norbery, E, Bery, P and Pedersen, J 2007. Genetic parameters for carcass traits and in vivo measured muscled and fat depth in Danish Texel and Shropshire. *Act Agriculture Scandinavica, Section A – Animal Science*, 57(2), 49-54.

McGowan E, Coffey M, Simm G and Mrode R 2022. Modelling growth in Suffolk and Charollais sheep populations using random regression models and validation of constrained polynomial correlation values. *Animal*. 17(5):100792.

McHugh, N, McDermott, K, Bohan, A, Farrell, L, Herron, J, Rabiou, T. 2022. Validation of maternal and terminal sheep breeding objectives using Irish field data, *Translational Animal Science*, Volume 6, Issue 3, July 2022

McLaren, A, Brotherstone, S, Lambe, NR, Conington, J, Mrode, R and Bunger, L 2015. The effects of different farm environments on the performance of Texel sheep. *Animal* 9(10), 1624-164.

- McLaren, A, Lambe, N, Morgan-Davies, C, Mrode, R, Brotherstone, S, Conington, J, Morgan Davies, J and Bunger, L 2014. Characterization of terminal sire sheep farm systems based on a range of environmental factors: a case study in the context of genotype by environment interactions using Charollais lambs. *Animal* 8, 867–876.
- McLaren A, Lambe NR, Brotherstone S, Conington J, Mrode R and Bunger L 2012a. Investigation into the presence of genotype by environment interactions ( $G \times E$ ) in Scottish Blackface lamb traits. *Small Ruminant Research* 105, 46–52.
- McLaren A, Lambe NR, Brotherstone S, Conington J, Mrode R and Bunger L 2012b. Do genotype by environment interactions ( $G \times E$ ) exist in Scottish Blackface ewes? *Proceedings of the British Society of Animal Science*, Nottingham, UK, p. 136.
- Meyer, K 2002 Estimates of covariance functions for growth of Australian beef cattle from a large set of data. CD-ROM communication No. 11-01 in *Proc 7th World Congr. Genet. Appl. Livest. Prod.*, Montpellier, France.
- Meyer, K 2001 Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning. *Genet. Sel. Evol.* 33:487–514
- Meyer, K 1999. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. *Journal of Animal Breeding and Genetics* 116, 181–205.
- Meyer, K. 1997. Estimates of genetic parameters for weaning weight of beef cattle accounting for direct maternal environmental covariances. *Liv. Prod. Sci.* 52, 187-199.
- Molina, A, Menéndez-Buxadera, A, Valera, M, Serradilla, JM 2007. Random regression model of growth during the first three months of age in Spanish Merino sheep. *Journal of Animal Science* 85, 2830-2839
- Mrode, R, Ojango, J, Ekine-Dzivenu, C, Aliloo, H, Gibson, J and Okeyo, MA 2021. Genomic prediction of crossbred dairy cattle in Tanzania: A route to productivity gains in smallholder dairy systems. *Journal of Dairy Science* 104, 11779-11789.
- Mrode, R, 2005. *Linear Models for the Prediction of Animal Breeding Values*, 2<sup>nd</sup> edition. CABI Publishing, Wallingford, UK.
- Mutch, E, Mrode, R, Coffey, M and Simm, G 2022. Modelling growth in Suffolk sheep with insufficient mature weight records for polynomial random regression models. Poster presented at 78<sup>th</sup> Annual British Society of Animal Science Conference, 12 -14 April 2022, Nottingham, United Kingdom.
- Nobre, P. R. C., Misztal, I, Tsuruta, S, Bertrand, J K, Silva, L O C and Lopes, PS 2003<sup>a</sup> Analysis of growth curves of Nellore cattle by multiple-trait and random regression models. *J. Anim. Sci.* 81:918–926

Osoro K, Olivan M, Celaya R and Martinez A 1999. Effects of genotype on the performance and intake characteristics of sheep grazing contrasting hill vegetation communities. *Animal Science* 69, 419–426.

Pollot, GE and Greeff, JC 2004. Genotype x environment interactions and genetic parameters for fecal egg count and production traits of Merino sheep. *Journal of Animal Science* 82, 2840-2851.

Prache, S, Schreurs, N and Guillier, L 2022. Review: Factors affecting sheep carcass and meat quality attributes. *Animal* 16, Supplement 1.

Pryce, J E, Arias, J, Bowman, PJ, Davis, SR, MacDonald, KA, Waghorn, GC, Wales, WJ, Williams, YJ, Spelman, RJ and Hayes, BJ 2012. Accuracy of genomic predictions of residual feed intake and 250-day body weight in growing heifers using 625,000 single nucleotide polymorphism markers. *Journal of Dairy Science* 95, 2108–2119.

Ptak, E and Schaeffer, LR 1993 Use of test day yields for genetic evaluation of dairy sires and cows. *Livest. Prod. Sci.* 34, 23–34

Ravagnolo, O and Misztal, I 2002. Studies on genetics of heat tolerance in dairy cattle with reduced weather information via cluster analysis. *Journal of Dairy Science* 85, 1586-1589.

Reddy, BVS, Kumari, PB, Sudhakar, K, Rao S, J and Pushpa RN, R 2018. Comparison of sire evaluation methods for growth traits in Nellore Jodipi sheep by LS, BLUP and DFREML methods. *J of Ent and Zoo Studies* 7 (1) 1161-1166.

Robinson, DL, 1996. Models which might explain negative correlations between direct and maternal genetic effects. *Liv. Prod. Sci.*, 45, 111- 122.

Roeber, F., Jex, A.R. & Gasser, R.B. 2013 Impact of gastrointestinal parasitic nematodes of sheep, and the role of advanced molecular tools for exploring epidemiology and drug resistance - an Australian perspective. *Parasites Vectors* 6, 153.

RStudio Team 2012. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.

Safari, E., Gilmour, A.R. and Fogarty, N.M. 2005 A review of genetic parameter estimates for wool, growth, meat and reproduction traits in sheep. *Livestock Production Science*, 92, 271-289

Sallam, AM, Ibrahim, AH and Alsheikh, S 2019. Estimation of genetic parameters and variance components of pre-weaning growth traits in Barki lambs. *Small Ruminant Research* 173, 94-100.

Santana ML Jr, Bignardi AB, Eler JP, Cardoso FF, Ferraz JBS 2013. Genotype by environment interaction and model comparison for growth traits of Santa Ines sheep. *Journal of Animal Breeding and Genetics* 130, 394-403.

Sargison, N. 2009 *Sheep Flock Health: A Planned Approach*, Oxford, UK: Blackwell Publishing

Sarti, F M, Lasagna, E, Giontella, A, Panella, F and Pieramati, C 2015. The use of a random regression model on the estimation of genetic parameters for weight at performance test in Appenninica sheep breed. *Italian Journal of Animal Science* 14, 310-315.

Schaeffer, LR, 2004. Application of random regression models in animal breeding. *Livestock Production Science* 86, 35-45.

Schenkel, FS, Schaeffer, LR and Boettcher, P 2002. Comparison between estimation of breeding value and fixed effects using Bayesian and empirical BLUP estimation under selection on parents and missing pedigree information. *Genetics Selection Evolution* 34, 41-59.

Schnyder, U, Hofer, A, Labroue, F, and Künzi, N 2001. Genetic parameters of a random regression model for daily feed intake of performance tested French Landrace and Large White growing pigs. *Genetics Selection Evolution* 33, 635–658.

Signet Breeding Services, 2021. Maternal and Terminal Sire Index Retrieved 1 Feb 2025, <https://signetdata.com/technical/genetic-notes/terminal-sire->



Webb, EC and Casey, NH 2010 Physiological limits to growth and the related effects on meat quality. *Livest Sci.* 130:33–40.

Woolaston, R.R. 1987 Genotype x environment interactions and their possible impact on breeding programs. In *Merino Improvement Programs in Australia*, ed B.J. McGuirk, Australian Wool Corporation, pp. 421-425

World Population Clock, 2022. worldometer current world population. Retrieved on 8 September 2022, from <https://www.worldometers.info/world-population/>.

Zeljic, K, Stanojevic, D, Bogdanovic, V, Mekic, C, Ruzic-Muslic, D, Maksimovic, N and Stojiljkovic, N 2019. Estimation of growth traits heritability coefficients in sheep. *Biotechnology in Animal Husbandry* 35, 347-355.

Zhang, Z, Kargo, M, Liu, A et al. 2019 Genotype-by-environment interaction of fertility traits in Danish Holstein cattle using a single-step genomic reaction norm model. *Heredity* 123, 202–214.