

**The Analysis and Use of Genotype by Environment Interactions
in Genetic Evaluations of Livestock and Plants**

Dominic Luke Waters

Bachelor of Science (Hons) University of New England

A thesis submitted for the degree of Doctor of Philosophy of the
University of New England

June 2023

School of Environmental and Rural Science

Faculty of Science, Agriculture, Business and Law

Abstract

This thesis explores methods for estimating genotype by environment ($G \times E$) interactions in livestock and plants. Genotype by environment interactions occur when the genetic architecture of a trait changes depending on the environment it exists in. They are particularly interesting as a source of genetic variation that could be utilised in breeding programs to select genotypes who have genetic merit that is more robust to environmental variation. This thesis aims to estimate genotype by environment interactions in livestock and plant populations using different methods and improve our understanding of how these interactions could be used in breeding programs to increase the robustness of agricultural populations to environmental variation.

The first experiment of this thesis investigated genotype by environment interactions in the bodyweight of Australian sheep using reaction norm and multi-trait models in combination with genomic data. It found significant variation in the slope of the reaction norm model that could be used to increase robustness of sheep, and that this variation was highly polygenic. It highlighted that both heterogenous genetic variance (scale-type $G \times E$) and heterogenous genetic correlations (rank-type $G \times E$) contributed to the variation in the reaction norm slope and found it could be important to separate these sources to better understand the genetic variation in robustness.

The second experiment of this thesis utilised a multi-environment trial of a Barley population to examine the effectiveness of two methods to partition the different types of $G \times E$ interactions when estimating the robustness of genotypes in reaction norm models. It found that genetic regression, which made breeding values for the slope independent of the intercept, was very effective in removing the impact of $G \times E$ interactions due to scale and isolating the variability across environments due to heterogenous genetic correlations. This enables the change in

genetic architecture of traits across environments to be studied more clearly in reaction norm models. We also showed that factor analytic models, which are an alternative to reaction norms, are better equipped to capture complex $G \times E$ interactions because of their flexibility.

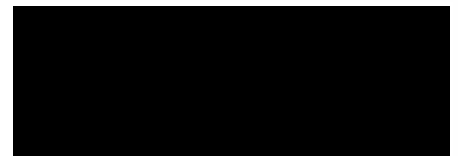
The third experiment of this thesis examined the use of factor analytic models to capture genotype by environment interactions in a multi-environment trial of sheep. The factor analytic models were able to approximate the unstructured genetic co(variance) matrix between 31 discrete environments using 85% fewer parameters than what would have been required with a multi-trait model. The model enabled the flock-years to be clustered by their similarity and showed that $G \times E$ interactions were large both between flocks and across years within flocks. It was unclear whether factor analytic models were preferable to reaction norms based on the goodness-of-fit tests, and the estimates of heritability, genetic variance and genetic correlations between environments were inconsistent between the models.

The final experimental chapter assessed the capacity of reaction norm models to predict the robustness of a sire's progeny performance across different growth environments. Using data collected in a research flock, the reaction norm models were predictive of the ability of a sire's progeny to reliably gain weight across different growth environments at an accuracy that was consistent with the power of the data. It then found that these breeding values for robustness were consistent with breeding values estimated using data from the wider industry population recorded by commercial stud breeders. Selection based on reaction norm breeding values could be used to increase the robustness of body weight gain in Australian sheep to variation in growth environments.

Declaration

I certify that the substance of this thesis has not already been submitted for any degree and is not currently being submitted for any other degree or qualification.

I also certify that any help received in preparing this thesis and all sources used have been acknowledged in this thesis.

A solid black rectangular box used to redact the signature of the author.

Dominic Luke Waters

Acknowledgements

I would like to thank my supervisors, Julius van der Werf and Sam Clark, for their guidance, encouragement and support. Together, your enthusiasm and knowledge has been invaluable in creating this thesis but also in developing my understanding of scientific research from an applied and theoretical perspective. Thanks also to all the other fellow scientists who contributed to this thesis.

This thesis would not have been possible without the support of my family and friends. Mum and Dad – thank you for your stable and unwavering support. Among so many other things, you have instilled me with confidence and the belief in my own ability, for which I am so grateful. Thanks also to the fishbowl genetics gang for allowing me to exercise some social skills outside of staring at a computer.

Finally, thank you Dannielle. I am so grateful for your constant love, patience and understanding.

I would also like to acknowledge the financial support provided by the Research and Training Program Scholarship and the Destination Australia Program Scholarship.

Table of Contents

Abstract	2
Declaration.....	4
Acknowledgements.....	5
Table of Contents.....	6
List of Figures	10
List of Tables	12
Chapter 1. Introduction	14
Chapter 2. Review of literature	19
2.1 Definition of genotype and environment	20
2.2 Genotype by environment interactions	23
2.3 Models for the analysis of genotype by environment interaction.....	24
2.3.1 Interaction model	25
2.3.2 Multi-trait model.....	27
2.3.3 Reaction norm model.....	29
2.3.4 Factor-analytic model	31
2.4 Robustness	34
Chapter 3. Genomic analysis of the slope of the reaction norm for body weight in Australian sheep	36
3.1 Abstract.....	37
3.2 Background	38
3.3 Methods.....	40
3.3.1 Data structure	40
3.3.2 Preparation of the reaction norm data.....	41
3.3.3 Genotype data	43

3.3.4	Reaction norm analysis	44
3.3.5	Genome-wide association (GWA) analysis	46
3.4	Results.....	47
3.4.1	Reaction norm models	47
3.4.2	GWA analysis and scale correction	50
3.5	Discussion	54
3.5.1	Reaction norm models	54
3.5.2	Genome-wide association analysis	56
3.6	Conclusions.....	59
Chapter 4. Partitioning the forms of genotype by environment interaction in the reaction norm analysis of stability		62
4.1	Abstract	63
4.2	Introduction.....	63
4.3	Methods.....	66
4.3.1	Phenotypic data.....	66
4.3.2	SNP data.....	67
4.3.3	Analysis.....	67
4.3.4	Reaction norm models	68
4.3.5	RN overall breeding value and stability estimates.....	70
4.3.6	Factor analytic models	72
4.3.7	Overall breeding value and stability in FA models	74
4.4	Results.....	77
4.4.1	Model summary	77
4.4.2	Magnitude of $G \times E$	79
4.4.3	Loadings of factor analytic models.....	80
4.4.4	Comparison of stability measures.....	81
Chapter 5 Factor analytic models applied to the analysis of genotype by environment interactions and robustness in livestock.....		93

5.1 Abstract	94
5.2 Background	95
5.3 Materials and Methods	97
5.3.1 Dataset	97
5.3.2 Factor analytic models	101
5.3.3 Post analysis of factor-analytic models	102
5.3.4 Reaction norm models	104
5.4 Results	106
5.4.1 Model fit	106
5.4.2 G×E interaction analysis with FA2	108
5.4.3 Comparison with reaction norm models	111
5.5 Discussion	116
5.6 Conclusions	119
Chapter 6 Validation of reaction norm breeding values for robustness in Australian sheep	122
6.1 Abstract	123
6.2 Background	123
6.3 Materials and Methods	126
6.3.1 Data	126
Part 1. Cross validation and forward prediction within the INF/RF dataset	126
6.3.2 Cross-validation	128
6.3.3 Forward prediction	130
6.3.4 Realised performance	131
Part 2. Validation of robustness EBVs between datasets	131
6.3.5 Data	132
6.3.6 Reaction norms	134
6.6 Results	136
Part 1. Cross validation and forward prediction within the INF/RF dataset	136

Part 2. Validation across datasets.....	137
6.7 Discussion.....	139
6.8 Conclusions.....	143
Chapter 7. General Discussion.....	146
7.1.1 The statistical and biological definition of robustness.....	149
7.1.2 The choice of model for estimating $G \times E$ interactions and robustness	154
7.1.3 The potential of selection for robustness in practice	158
7.2 Conclusions.....	162
Appendix A.....	164
Appendix B	168
Chapter 8. Consolidated Reference List	171

List of Figures

Chapter 3

Figure 3. 1 Estimates of genetic variance and heritability along the environmental gradient.	49
Figure 3. 2 Estimates of genetic correlations between low- average-, and high- growth environments.....	50
Figure 3. 3 Scatterplots and correlations between SNP effects for the direct slope and scale-corrected slope against the intercept.....	51
Figure 3. 4 Reaction norms before and after scale-correction.....	52
Figure 3. 5 Genome-wide SNP associations for the slope using direct and scale-corrected GEBV for the slope.....	53

Chapter 4

Figure 4. 1 Schematic example of the regression of four genotypes (coloured lines) across a common factor, 'x' in a FA model.....	76
Figure 4. 2 Additive genetic variance across the environmental covariable for the four models.....	79
Figure 4. 3 Scatterplots between stability GEBVs from RN models compared with GEBVs for responsiveness to the second common factor in FA-2 and RR-2.	82
Figure 4. 4 The relationships between overall performance and stability within the RN and FA models.....	84
Figure 4. 5 Reaction norms before and after applying the scale corrections.....	86

Chapter 5

Figure 5. 1 The number of parents (sires, dams, grandsires and grandams) in common between the flock-years	99
Figure 5. 2 Distribution of the percentage of genetic variance accounted for by the common factors in the FA models across the different environments.....	107
Figure 5. 3 Distribution of the rotated loadings for the first and second common factor across environments in FA2.....	109

Figure 5. 4 EBVs for overall performance and responsiveness to the second common factor for 351 sires with at least 25 progeny.	109
Figure 5. 5 A heatmap showing the genetic correlations between the 31 different environments in the data..	110
Figure 5. 6 Distribution of heritability across the environments in FA2.	111
Figure 5. 7 The relationship between the average EC value within each flock-year and the loadings for the first and second common factor in FA2.....	112
Figure 5. 8 The relationship between estimates of genetic variance and heritability for each flock-year, as estimated in FA2 and RN_CG_FY.	113
Figure 5. 9 Genetic correlations between all flock-years as estimated in the reaction norm model and factor analytic model.....	114
Figure 5. 10 Scatterplots of EBVs for 351 sires with more than 25 progeny.	115
Figure 5. 11 Scatterplots of the absolute value of reaction norm and factor analytic robustness EBVs for 351 sires with more than 25 progeny.....	115

Chapter 6

Figure 6. 1 Distribution of 18,171 lambs across the mean-centred EC.	128
Figure 6. 2 Distribution of lambs in 2007-2010 and 2011 across the mean-centred EC.....	130
Figure 6. 3 Distribution of lambs across the EC in the MERINOSELECT and INF/RF datasets.....	133
Figure 6. 4 Genetic variance and heritability across the EC in the MERINOSELECT and INF/RF datasets.	138

Chapter 7

Figure 7. 1 Reaction norm for 5 genotypes where the interaction across the environmental covariable, E, is only due to scale-type $G \times E$ interaction.	151
--	-----

List of Tables

Chapter 3

Table 3. 1 Variance of the additive genetic intercept and slope, the residual intercept, slope and covariance, as well as genetic group variance. The correlation between additive genetic intercept and slope and the log-likelihood are also reported48

Table 3. 2 SNPs associated with level of robustness54

Table 3. 3 List of known genes in a 500-kb window surrounding the most significant SNP..54

Chapter 4

Table 4. 1 Summary of model fit.78

Table 4. 2 Pearson correlations between GEBVs for overall performance in the RN models and FA models79

Table 4. 3 Minimum, maximum and mean pairwise correlations between the 15 environments for the four models.....80

Table 4. 4 Absolute value of the mean REML estimates of the positive and negative loadings for the two common factors in FA-2 and RR-2.81

Chapter 5

Table 5. 1 Numerical summary of the dataset 100

Table 5. 2 Summary of model fit for the different models.. 107

Table 5. 3 Variance of the genetic intercept, slope, the ratio and the correlation between them 111

Chapter 6

Table 6. 1 Numerical summary of animals within each fold, as well as the progeny of sires not eligible for validation but used in the training data 129

Table 6. 2 Numerical summary of animals in the INF/RF and MERINOSELECT 133

Table 6. 3 Correlation between reaction norm EBVs for sires and the realised intercept and slope of the random regression of pre-corrected progeny phenotypes across the EC in the test sets..... 136

Table 6. 4 Variance of additive genetic intercept, slope and the correlation between them, as well as genetic group and permanent environmental dam variance for the reaction norms models	137
Table 6. 5 Variance of the additive genetic intercept, slope, the ratio and the correlation between them, as well as genetic group, permanent environmental dam, and maternal genetic variance for the reaction norms estimated in the INF/RF and MERINOSELECT (MS) data.	137
Table 6. 6 Correlation of reaction norm EBVs for intercept, slope and scale-corrected slope for sires.....	139

Chapter 1. Introduction

There are an infinitely large number of characteristics which we can measure on an organism. Some examples include the height of a person, the yield of a wheat variety, or the growth rate of a lamb. These characteristics, called phenotypes, are the sum of genetic (**G**) and environmental (**E**) factors, which are explained by the equation: $P = G + E$. The relative contribution of genetic and environmental factors depends on the phenotype that is measured. In some instances, a phenotype is controlled by a single gene, with little influence of environmental factors. The expression of fragrance in rice and the presence or absence of horns in sheep are examples of this (Bradbury et al., 2005; Duijvesteijn et al., 2018). More often, phenotypes are influenced by many genes with small individual effects, as well as many environmental factors. For instance, the weight of a lamb 300 days after birth (**P**) is influenced by its genetic composition (**G**), as there might be particular genes that increase appetite, or the conversion of feed into muscle and fat. It is also influenced by how much feed was available (**E**), as lambs with access to more calories will grow more. Genetic analysis is often undertaken with the assumption that the effect of **G** is independent of **E**. However, imagine that feed (**E**) is significantly limited, as can happen in a drought year. Genes that increase appetite would clearly not have the same effect on growth as when feed is unlimited. This dependency between **G** and **E** is an example of a genotype by environment (**G**×**E**) interaction, where the phenotype is described by: $P = G + E + G \times E$.

The practical consequence of **G**×**E** interactions in breeding programs is that the best genotype in one environment is not necessarily the best in another environment, which creates a dilemma when making selection decisions. A sensible response to this problem would be to ensure that the population has a consistent environment. By decreasing the variation in **E**, the **G**×**E** interaction will reduce, making selection more straightforward. The glasshouse production of horticultural crops such as tomatoes is an example of this approach (Heuvelink, 2005).

Environmental factors such as the composition of nutrient solutions, temperature, humidity, carbon dioxide concentration and light intensity can be manipulated to maintain the optimal environmental conditions for production, while the incidence of pests and diseases is greatly reduced. In livestock, the intensive farming of poultry also allows a similar level of environmental control. However, this is not possible in all types of agricultural systems.

There are far fewer opportunities to control environmental factors in extensive farming systems such as sheep and cattle grazing, or field cropping. The size of these operations prohibits the same level of environmental control. Consequently, the environment which animals and plants are exposed to can vary substantially between different years and locations, possibly leading to G×E interactions. When this occurs, the merit of the genotype carried by each animal or plant can change depending on the location or year it will exist in.

Location-based G×E interactions arise because of environmental factors that are specific to a particular location. Examples could include the soil quality, pasture type and management practices of a region. This type of G×E interaction can be managed by developing specific breeding programs for different locations (James, 1961). An example of this in livestock is the Interbull Analysis, which estimates the genetic merit of each genotype as dependent on the country it will be used in (Nilforooshan & Jorjani, 2022). However, this strategy is limited by the resolution in which a location can be defined in an analysis; for instance, there can be significant environmental variation within a country. More importantly, location-based G×E interactions also represent only a part of the G×E picture.

Yearly (or seasonal) environmental variations are another source G×E interactions. They are more difficult to manage in breeding programs, as the best genotype will change depending on the year it will be used in. To always use the best genotypes each year, one would need to know 1) the future environmental conditions (e.g., rainfall, temperature, disease burden), and 2) the

ranking of genotypes given the environmental conditions. Even if this information was known accurately, it would be unwise to select genotypes for the environmental conditions of a single year, as a herd of cows, a flock of ewes, or retained seed typically needs to be used for production across multiple years. When genotypes are likely used across a range of environments in which the effects are not predictable in advance, a more practical solution would be to select genotypes that have genetic merit that is more ‘robust’ to environmental variation. By increasing robustness, the impact of environmental variation on the performance on genotypes across years will be reduced. Genotypes with greater robustness are likely to become more valuable in the future, as environmental variability is predicted to increase with climate change (Powell et al., 2012; Rojas-Downing et al., 2017). Robust genotypes may also help to manage location-based G×E interactions, as they could be used across a wider geographical range of environments. However, an understanding of the methods available to capture G×E interactions and genetic variation in robustness is required for it to be utilised effectively. cross

A common way to model G×E interactions and estimate robustness is the reaction norm model. Reaction norm models use random regression to regress the breeding value of genotypes across an environmental covariable that describes the quality of the environment, such as temperature and humidity (Stefani et al., 2022), or the mean performance of a contemporary group (Carvalho et al., 2019). This allows G×E interactions due to the environmental covariable to be estimated, as well as the robustness of the genetic merit of individual genotypes across the environmental covariable. Although these models have been available for some time (Calus et al., 2002; Kolmodin et al., 2002), there are many livestock species for which genetic variation in the robustness of important traits has yet to be explored. There are even fewer examples of genomic data being used in conjunction with these models, which offers the potential to explore

the molecular framework that gives rise to G×E interactions and variation in the robustness of animals and plants (Li et al., 2018; Silva et al., 2014).

Factor analytic models have also been developed for modelling G×E interactions and robustness, although their application has been mostly limited to plants (Kelly et al., 2007). Like reaction norms, these models capture G×E interactions by regressing the breeding value of genotypes on environmental covariables; however, the covariables are estimated from the data, rather than being user-defined. While reaction norms typically only use a single covariable, factor analytic models can easily accommodate several covariables. The additional flexibility and unsupervised nature of factor analytic models could reveal novel insight into how G×E interactions are captured in livestock.

Although there has been a long-standing interest in the selection of both animals (Falconer, 1990; James, 1961) and plants (Finlay & Wilkinson, 1963; Perkins & Jinks, 1968; Yates & Cochran, 1938a) with genetic merit that is more robust to environmental variation, selection has been adopted more widely in the plant breeding context (Khanna et al., 2022; Mickelbart et al., 2015; Qian et al., 2017). There is only one example of selection for robustness directly being implemented in livestock (Nguyen et al., 2017). This is despite the existence of significant G×E interactions across a variety of livestock species (Hayes et al., 2016; James, 2009; Rauw & Gomez-Raya, 2015). This suggests that more research is needed to quantify the magnitude of genetic variation in robustness for species where G×E interactions are important and to increase our understanding how this genetic variation could be applied to increase robustness. The aim of this thesis was to investigate genotype by environment interactions and estimate genetic variation in robustness in both animal and plant populations, and furthermore, to improve our understanding how this variation could be used in breeding programs to increase robustness.

Chapter 2 is a review of literature that discusses the theory and importance of genotype by environment interactions in livestock and plants. It focuses on the different methods for estimating of $G \times E$ interactions and how measures of robustness can be derived from these models. It will also discuss how robustness has been defined in the literature.

Chapter 3 uses a reaction norm applied to the bodyweight of sheep to investigate genotype by environment interactions due to different growth environments and explore associations of genomic regions with the reaction norm slope. In addition, it examines how the different types of $G \times E$ interactions may affect the interpretation of the slope at a genomic level.

Chapter 4 examines how the different types of $G \times E$ can be partitioned from the reaction norm slope, leading to a better understanding of robustness. We analyse phenotypic and genomic data from an Australian Barley breeding program using factor-analytic models to show how reaction norm slope variation can be partitioned into the different types of $G \times E$.

Chapter 5 examines the application of factor-analytic models, which are routinely applied in the genetic analysis of plant data, to model $G \times E$ interactions in livestock data. We show how they can group environments in which performance is more correlated, and rank genotypes on overall performance and robustness. The models were also compared with RN models.

Chapter 6 examines the ability of reaction norm models to predict the robustness of post-weaning weight to different growth environments in Australian lambs. It discusses the practicality of implementing selection for robustness using these models.

Chapter 7 is the general discussion of this thesis. It discusses three main topics arising from this work. These are: 1) the statistical and biological definition of robustness, 2) the choice of model for estimating $G \times E$ interactions and robustness, and 3) the potential of selection for robustness in practice.

Chapter 2. Review of literature

Breeding programs aim to improve the performance of animals and plants for traits relevant to a breeding objective. This requires the identification and dissemination of genotypes with superior performance for the desired traits. However, the ranking of genotypes for performance can be dependent on the environment they will exist in because of genotype-by-environment ($G \times E$) interactions. When significant, these interactions need to be accommodated in the genetic evaluation models used to rank genotypes for performance.

In livestock evaluations, phenotypic variance due to $G \times E$ interaction is often modelled as a nuisance parameter, such as a sire by flock-year interaction term. This discards the variation due to $G \times E$ interactions and yields a ranking of genotypes based on performance in the average environment (Li et al., 2015). However, the variance due to $G \times E$ interactions might contain useful information on how the performance of individual genotypes changes in different environments. This information could be particularly valuable in extensive production systems such as sheep and cattle grazing, where genotypes need to perform across highly variable environments (Derner et al., 2017). Hence, it is important to explore methods which model $G \times E$ interactions and capture individual responses to different environments.

It is relatively more common for routine genetic evaluations of field crops to explicitly parameterise $G \times E$ interactions and capture variation in robustness (van Eeuwijk et al., 2016). The models applied in this space enable the identification and propagation of high-yielding varieties with robust performance across different environments (Malosetti et al., 2013). The knowledge available from modelling $G \times E$ interactions in plants could be beneficial for understanding its use in livestock data.

This literature review will explore the theory of $G \times E$ interactions, and the importance of these interactions in animal and plant populations. In particular, it will focus on the different methods

for estimating of G×E interactions, and how measures of robustness can be derived from these models. It will also briefly explore how robustness has been defined in the literature.

2.1 Definition of genotype and environment

An analysis of G×E interaction requires at least two genotypes that are expressed in at least two environments. It is therefore important to first define the terms genotype and environment. The genotype is a unit that is based on the research question and available data. Most broadly, genotypes can consist of groups of individuals, such as different sub-species (Bressan et al., 2011) or breeds in livestock (Brown et al., 1993; Vercoe & Frisch, 1992), and different landraces in plants (Zeven, 1998). In this case, the interest is in contrasting the merit of groups of individuals under certain environmental conditions. Genotypes can also be defined as an individual within a group, such as a sire (Calus et al., 2002), the parents used in a cross (Cullis et al., 2014), or more commonly, the individual animal or plant variety on which the phenotype is measured (Madsen et al., 2018; Smith et al., 2001). Here, the interest is in the ranking of individuals for selective breeding. However, it could be more instructive to think of a genotype in terms of the genes that an individual (or group of individuals) carries. In most cases, the phenotype of a single animal or plant will only be recorded in one environment. Therefore, G×E interactions cannot be estimated based on the single plant or animal alone, but instead on the genes that it shares with other individuals that exist in different environments based on either pedigree or genomic information (Hayes et al., 2016). Hence, all G×E interactions are implicitly estimated at the level of the genes, or the ‘genotype’, that an individual or group of individuals carries. This view is explicitly realised with the use of single nucleotide polymorphism (SNP) markers, which allow G×E interactions to be explored at the genomic level (Silva et al., 2014; Tolhurst et al., 2019).

The definition of environment is less straightforward but is still determined by the research question and available data. Environments refer to an identifiable ‘macro-environmental’ factor that varies between genotypes (Mulder et al., 2013). Such factors can consist of directly measured environmental variables that are recorded during the life of an animal or plant, such as temperature (Hedhly et al., 2005), sunlight (Hébert et al., 2001), nutrition (Horan et al., 2005) or a combination of different variables (Jarquín et al., 2014; Ly et al., 2018; Nel et al., 2023). Some studies in dairy have used a temperature-humidity index to capture G×E interactions in daily milk measurements with success (Cheruiyot et al., 2020; Haile-Mariam et al., 2008). However, many of the phenotypes relevant to extensive agricultural production (e.g., live weight, wool quality, grain yield) are developed over a long period in which many environmental variables interact and cumulatively impact to the resulting phenotype (Basford & Cooper, 1998; Monaghan, 2008). This complexity can make it difficult to relate phenotypic variation to the individual environmental variables that preceded it, and the interpretation can be confounded by important variables that might not have been measured. This is further compounded by difficulties in obtaining accurate environmental measurements in extensive farming operations (Bahlo et al., 2019). An example in sheep used clustering based on temperature and normalised difference vegetation index (NDVI) data measured from satellites to form grazing environments (Whelan et al., 2009). While the creation of grazing environments offered the potential to understand the underlying components of the G×E interactions, they provided a poorer fit compared to a simpler sire by flock interaction model. An alternative to field-based studies is to run an experiment, in which environmental variables are carefully controlled and contrasted (Horan et al., 2005; Seixas et al., 2017; Zhao et al., 2004). However, the size of experimental trials limits their statistical power relative to large datasets from field studies. Experience in plant breeding has also found that G×E interactions under experimental conditions are often difficult to translate to the field environment (Telfer

et al., 2022; van der Putten et al., 2016). In response, the integration of measured environmental variables to model G×E interactions in field studies is a particularly active field of research, with several new models being developed to handle this data more efficiently in plants (Cossa et al., 2022; Tolhurst et al., 2022).

An alternative to using measured environmental data is to form groups of environments based on geographical regions (Carabaño et al., 1990). It is reasonable to expect that weather conditions, farming practices and other environmental qualities differ by region. It is also an attractive approach for commercial evaluations, as only the location is required for analysis. The global evaluation of dairy bulls is an example of this, where countries are treated as different environments (Nilforooshan & Jorjani, 2022). Similarly, the relative performance of plant varieties in Australia is reported by the National Variety Trials (NVT) for specific regions within Australia (Eichi et al., 2020). Environments can also be based on the type of farming practices, such as automatic and conventional milking systems in dairy cattle (Mulder et al., 2004), irrigated and rainfed yield in wheat (Chaudhary et al., 2022) and the overall intensity of production systems (Wahinya et al., 2020). These definitions can help answer practical questions, such as whether breeding values estimated in a particular region or production system are valid in another. However, a weakness of this approach is that environmental conditions can vary considerably within a region or production system, which could make predictions of merit within an environment unreliable. This could be handled by defining environments at the farm-level, such as a flock, herd or trial location (Jang et al., 2022). Variation between years within a farm can also be captured by considering each farm-year combination as an environment (Smith et al., 2001). However, this results in a large number of discrete environments which can be difficult to model informatively.

Another option is to define the environment for groups of individuals (such as a contemporary group or trial) using their mean performance as a quantitative proxy for the quality of the

environment. Examples include milk production (Calus & Veerkamp, 2003), the yield of a trial (Tan et al., 2020; Yates & Cochran, 1938), or the growth rate of an animal (Cardoso & Tempelman, 2012; Carvalheiro et al., 2019; Li & Hermes, 2016). This allows environments to be described continuously, which avoids the loss of environmental information associated with categories. It also indexes the cumulative effect of important environmental variables into a single value, without the need to directly measure them. However, the mean performance can also be affected by other factors, such as genetic merit, breed, sex, and other effects, which can bias the estimate of mean performance when groups are unbalanced (Su et al., 2006). This can be overcome by obtaining the effect of the environment on the group performance using best linear unbiased estimation (BLUE) to adjust for these factors (Carvalheiro et al., 2019; Daltro et al., 2022; Oliveira et al., 2018; Santana et al., 2015). Unlike measured environmental variables, the mean performance is estimated and therefore subject to error. Simulations have shown that a Bayesian approach where the mean performance is estimated jointly within the reaction norm model accounts for this uncertainty optimally (Su et al., 2006). This method can have very long computational times. An alternative is to use large contemporary groups with high levels genetic linkage to maximise the accuracy in which mean performance is estimated (Calus et al., 2004).

2.2 Genotype by environment interactions

Genotype by environment interactions occur when the effect of a genotype depends on the environment it exists in. There are two types of $G \times E$ interaction: 1) rank-type, which occurs because of genetic correlations between environments that are less than one, and 2) scale-type, which occurs because of heterogeneous genetic variance across environments (James, 2009). Rank-type $G \times E$ interactions are the most important, as it means that the ranking of genotypes changes between environments. As a result, optimal selection in one environment will be sub-optimal in another. The magnitude of rank-type $G \times E$ interaction is assessed by estimating the

genetic correlation between the same trait expressed in two different environments, with a smaller correlation indicating greater re-ranking. In practice, a genetic correlation less than approximately 0.80-0.60 is significant enough to justify separate breeding programs for two different environments (James, 1961; Mulder et al., 2006; Robertson, 1959). Correlations of this magnitude imply that performance in the different environments can be viewed as separate traits with different genetic architectures. However, any correlation less than 1 indicates there is genetic variation between genotypes in how they perform across different environments. This genetic variation could be used to select genotypes with genetic merit that is more ‘robust’ to environmental variation (Mulder, 2017).

Scale-type $G \times E$ interactions arise when the genetic variance between environments is different. While this causes the effect of a genotype to change in magnitude, it does not affect the genetic correlation and therefore ranking of genotypes between environments. However, scale-type $G \times E$ interactions can cause re-ranking of genotypes based on a selection index of multiple traits (Namkoong, 1985) and can also bias the estimation of breeding values. These problems can be managed by accommodating heterogeneous variances in the model of analysis or directly transforming the data (Meuwissen et al., 1996; Wiggans & VanRaden, 1991).

2.3 Models for the analysis of genotype by environment interaction

To estimate genotype by environment interactions, the performance of a genotype needs to be recorded in different environments. This is relatively easy to achieve in fully inbred lines or first-cross hybrids of plants, as we can directly assess the performance of identical copies of genotypes in different environments. In contrast, populations of livestock and out-crossing species of plants consist of genetically distinct individuals who only perform in one environment. Within these populations, however, there are related individuals that share fractions of their genome. We can use relationships between individuals in different

environments, estimated using either pedigree or genomic data, to capture G×E interactions in statistical models. Genetic linkage between different environments is created by assigning parents to have progeny in multiple environments (Falconer & Mackay, 1996; Kerr et al., 2002; Kolmodin et al., 2002), although the use of genomic data means this structure might no longer be as critical (Hayes et al., 2016). With genomic data, it is the individual SNP genotype that needs to be replicated in different environments, rather than progeny. This is leveraged in genomic-based G×E interaction studies of humans, where highly related individuals are actually excluded from the analysis (Ni et al., 2019). Genomic data also enables G×E interactions can be estimated more accurately by increasing the genetic linkage between different environments (Mulder, 2016).

Provided there is genetic linkage across environments, there are a range of models which can be used to estimate G×E interactions. All of the models extend from the baseline animal model, which ignores G×E interactions:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a} + \mathbf{e} \quad (1)$$

Where $\mathbf{y}$ is a vector of phenotypes, $\mathbf{X}$ is an incidence matrix for the fixed effects $\mathbf{b}$, $\mathbf{Z}_1$ is an incidence matrix relating the phenotypes in $\mathbf{y}$ to the breeding values in $\mathbf{a}$, and $\mathbf{e}$ is the residual. The model estimates only one breeding value per genotype, and the distribution of the breeding values are assumed to follow: $[\mathbf{a}] \sim N(0, \mathbf{A} \otimes \sigma_a^2)$, where $\mathbf{A}$ is a relationship matrix calculated by either pedigree (Henderson, 1976) or genomic (VanRaden, 2008; Yang et al., 2010) information and σ_a^2 is the additive genetic variance.

2.3.1 Interaction model

The simplest approach to modelling G×E interactions is to add an additional genotype by environment interaction variable to the animal model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{s} + \mathbf{e} \quad (2)$$

The model is identical to Eq.1, except for $\mathbf{Z}_2$, which is an incidence matrix relating the phenotypes in $\mathbf{y}$ to the G×E interactions interaction effects in $\mathbf{s}$. The elements of $\mathbf{s}$ refer to a genotype × environment combination, and the effects capture the deviation from the expected performance of the genotype in the specific environment. The genotypes in $\mathbf{s}$ need to have records in multiple environments, such as a sire with progeny recorded in different environments, or a plant variety recorded in multiple trials. The variance of the interactions effects, $\mathbf{s}$, is given by $[\mathbf{s}] \sim N(0, \mathbf{I} \otimes \sigma_s^2)$, where $\mathbf{I}$ is an incidence matrix. The size of the interaction variance (σ_s^2) relative to the genetic or overall phenotypic variance indicates the magnitude of the G×E interactions. By including the interaction variance in the model, variation in $\mathbf{y}$ that is not due to the main effect of σ_a^2 is partitioned and treated as a residual effect. The resulting breeding values represent performance in the average environment.

The interaction model is commonly used in large genetic evaluations where G×E interactions are important, such as in Australian sheep (Li et al., 2015) and beef cattle (Graser et al., 2005), and American beef cattle (Golden et al., 2009). The model was also used for plant breeding evaluations but has since been replaced by more complex models (Malosetti et al., 2013; Smith et al., 2001). The environment is usually represented as a location (e.g., herd, flock, farm) or a year × location combination, and the interaction usually explains between 0 to 10% of the phenotypic variance in livestock (Jang et al., 2022; Jeyaruban et al., 2009; Li et al., 2015; Pollott & Greeff, 2004; Whelan et al., 2009). The interaction model is appealing for the large routine evaluations of livestock as it allows many environments to be included for a relatively low computational burden. However, the simplicity also reduces the accuracy and interpretability of the estimated G×E interactions. In livestock, the interactions are estimated at the level of the sire, which only represents a fraction of the genetic relationships that exist between different environments. The genotypes are assumed to be independent when estimating the interaction variance, which is not true in practice; genotypes (e.g., sires) are typically

related. The model implicitly fits a uniform structure between environments (Smith et al., 2005), which assumes that the genetic variance and genetic covariance is constant between environments. Consequently, scale-type G×E interactions are not accommodated, and the level of rank-type G×E between all pairs of environments is assumed to be constant. The resulting breeding values also do not contain information on how individual genotypes respond to the environments. Therefore, the interaction model is less commonly applied in research settings, in favour of more informative approaches which allow the genetic variance and correlation to vary between environments.

2.3.2 Multi-trait model

G×E interactions can also be modelled by considering performance in different environments as separate but potentially correlated traits in a multi-trait model (Falconer, 1952). In the case of only two environments, the multi-trait model is described as:

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1 & 0 \\ 0 & \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{z}_1 & 0 \\ 0 & \mathbf{z}_2 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} \quad (3)$$

The model is identical to Eq.1, except that the components have been divided into two sections based on the environment that the phenotype was measured in, denoted by the subscripts (1 or 2). The resulting estimates of the fixed effects, breeding values and residual variance are specific to each environment. The breeding values are assumed to follow a bivariate normal

distribution: $\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix} \sim N(0, \mathbf{A} \otimes \mathbf{K})$, where $\mathbf{K} = \begin{bmatrix} \sigma_{a_1}^2 & \sigma_{a_1 a_2} \\ \sigma_{a_2 a_1} & \sigma_{a_2}^2 \end{bmatrix}$ and $\mathbf{A}$ was the same as in Eq.1.

Note that the multi-trait model estimates two breeding values for each genotype (one for each environment) and captures the extent of scale-type ($\sigma_{a_1}^2$ vs $\sigma_{a_2}^2$) and rank-type (genetic correlation = $\frac{\sigma_{a_2 a_1}}{\sqrt{\sigma_{a_1}^2 \times \sigma_{a_2}^2}}$) G×E interactions between environments. The residual variances ($\mathbf{e}_1$

and $\mathbf{e}_2$) are estimated independently for each environment (i.e., zero covariance), as individuals can only exist in one environment.

The multi-trait model is particularly useful for modelling $G \times E$ interactions between a small number of discrete environments. For example, Ducrocq et al. (2022) used a four-trait model to explore $G \times E$ interactions in Holstein cattle between four regions in South Africa, focusing on production traits and age at first calving. The genetic correlations ranged from 0.62 to 0.87, and were all statistically different from 1, indicating that genotypes re-ranked between environments. Genetic variance was also heterogeneous across the environments, suggesting the importance of scale-type interactions. Another example in beef cattle estimated $G \times E$ interactions in growth between low and high-altitude farms (Williams et al., 2012). The genetic correlation between low and high altitude was 0.74, meaning that the genetic architecture for growth depended on the altitude it was measured at. There are many other examples of the multi-trait approach being used to answer similar questions in beef and dairy cattle (Fikse et al., 2003; Haile-Mariam et al., 2008; Jeyaruban et al., 2009; Saavedra-Jiménez et al., 2013), sheep (Hazard et al., 2016; Hollema et al., 2018), and pigs (Gourdine et al., 2019; Wallenbeck et al., 2009), as well as various species in plant breeding (Raman et al., 2023; Telfer et al., 2021). However, most of these studies were limited to a small number of environments due to computational constraints. The number of parameters required to model the genetic covariance between environments increases exponentially with the number of environments (p), according to $p(p + 1)/2$. This can be managed by absorbing environments into larger groups (Calus et al., 2002; Hollema et al., 2018; Wahinya et al., 2020), but this comes at the loss of environmental information within the groups. Another approach is to use many pairwise bivariate analyses between environments (Ismael et al., 2016; Li & Hermes, 2013; Saavedra-Jiménez et al., 2013), but this increases the risk of obtaining parameter estimates that are biased or outside of the parameter space (Sae-Lim et al., 2016). When the number of environments is large, models

which estimate the genetic (co)variance between environments using a reduced number of parameters should be explored.

2.3.3 Reaction norm model

Reaction norm models can be used to model G×E interactions when the quality of an environment can be described by a continuous environmental covariable (EC). These models allow the breeding value of each genotype to change as a function of the EC using random regression. A linear function is most often used to model reaction norms, and it is fit as an extension to baseline animal model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a}_0 + \mathbf{Z}_2\mathbf{a}_1 + \mathbf{e} \quad (4)$$

Where the components $\mathbf{y}$, $\mathbf{X}$, $\mathbf{Z}_1$, and $\mathbf{b}$ are the same as in Eq.1, and $\mathbf{a}_0$ is also the same as $\mathbf{a}$ in Eq.1. The only difference is the matrix $\mathbf{Z}_2$ and the vector $\mathbf{a}_1$, where $\mathbf{Z}_2$ contains the values of the EC on the diagonal and links the records in $\mathbf{y}$ to the random regression coefficients (or breeding values) for the slope in $\mathbf{a}_1$. The variance of the regression coefficients are modelled

as: $\begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \end{bmatrix} \sim N(0, \mathbf{A} \otimes \mathbf{K})$, where $\mathbf{K} = \begin{bmatrix} \sigma_{a_0}^2 & \sigma_{a_1 a_0} \\ \sigma_{a_0 a_1} & \sigma_{a_1}^2 \end{bmatrix}$, and the genetic covariance between points

along the EC can be obtained as $\mathbf{G} = \boldsymbol{\Phi}\mathbf{K}\boldsymbol{\Phi}'$, where $\boldsymbol{\Phi}$ consists of two columns; the first is a vector of 1's, and the second is a vector of EC values for which genetic (co)variances are to be estimated. Therefore, scale-type and rank-type G×E interactions can be estimated between any point along the EC in the data. In addition to modelling G×E interaction, the regression coefficients (i.e., the slope) capture the change in breeding value for each genotype across the EC. This enables the identification of genotypes with genetic merit that is more robust to changes in the EC level (Kolmodin et al., 2002). Another advantage of reaction norms is that the number of parameters required to estimate the G×E interactions between environments depends on the order of the polynomial rather than the number of environments. For 100

environments, a multi trait model would need to estimate 5050 parameters to model the genetic (co)variance, whereas a linear reaction norm ($k = 2$) would require $k(k + 1) / 2$, or 3 parameters (an intercept, slope and a covariance between them). There are also fewer parameters needed to estimate the residual variance, as it can either be modelled independently for discrete sections of the EC (Calus & Veerkamp, 2003; Cardoso & Tempelman, 2012; Lillehammer et al., 2009), or modelled as a continuous function along the EC, similar to the genetic variance (Carvalho et al., 2019; Ni et al., 2019). However, the reduction in parameters is only possible by assuming that the patterns in genetic co(variance) and residual variance across environments are exclusively a function of the EC.

Reaction norms have been used extensively to model G×E interactions in dairy cattle (Calus et al., 2002, 2006; Calus & Veerkamp, 2003; Hayes et al., 2003, 2009; Mulder et al., 2004; Shi et al., 2021), where the genetic correlations between extreme levels of the EC (i.e., low vs high) tended to be around 0.80 or higher. Even though these correlations were generally high enough to not impact the efficiency of a breeding program, they still represented genetic variation between genotypes in their robustness to environmental variation. There are also several examples in beef cattle (Cardoso & Tempelman, 2012; Carvalho et al., 2019; Mota et al., 2020; Oliveira et al., 2018). Notably, Mota et al. (2020) estimated genetic correlations as low as -0.12 for scrotal circumference traits in Nellore cattle, suggesting the existence of large G×E interactions. There is often less data in extreme environments, so care should be exercised when reporting large interactions such as these.

Studies in sheep have also detected important G×E interactions and variation in robustness for different traits (Hollema et al., 2018; Pollott & Greeff, 2004; Santana et al., 2013). More recently, the direct inclusion of genomic data in reaction norm models has enabled the effect of individual SNPs on the reaction norm slope to be investigated (Silva et al., 2014). Most of these analyses have been in beef cattle (Carvalho et al., 2019; Mota et al., 2020; Oliveira et

al., 2018), while there are no studies utilising genomic reaction norm models in sheep, highlighting the potential for further research in this area.

In plant breeding, reaction norm models that regress on the mean performance of a trial are frequently referred to as “Finlay-Wilkinson regression” models (Finlay & Wilkinson, 1963). Although they have been used to estimate robustness in plants in recent studies (de Souza et al., 2020; Kusmec et al., 2017; Sandhu et al., 2021), they are being applied less frequently (van Eeuwijk et al., 2016). This is because it is uncommon for a single EC to fully explain the G×E interaction in a multi-environment trial (Kelly et al., 2007). Although reaction norms with multiple ECs can be fit (Calus et al., 2006), this requires the user to pre-select the ECs, which might still be uninformative. Instead, approaches that ‘find’ the ECs that describe the G×E interaction in the data are more popular, such as factor analytic models.

2.3.4 Factor-analytic model

When there are too many environments to fit a multi-trait model, and the G×E interactions are not sufficiently described by a reaction norm, factor analytic models can be used. These models approximate the unstructured genetic (co)variance matrix between environments by assuming it can be described by the regression of breeding values on latent variables, called common factors, which are essentially ECs that are estimated from the data (Meyer, 2009; Smith et al., 2005). Genetic variation that is unexplained by these common factors is captured by a residual genetic effect specific to each environment. These models were developed for the analysis of multi-environment trials of field crops, where each environment typically consists of a location × year combination (Smith et al., 2005). For a model with c common factors, the estimated breeding value (EBV) for genotype i in environment j is given by the multiple regression:

$$EBV_{ij} = \lambda_{a_{1j}}f_{a_{1i}} + \lambda_{a_{2j}}f_{a_{2i}} + \cdots + \lambda_{a_{cj}}f_{a_{ci}} + \delta_{a_{ij}} \quad (5)$$

where λ_{a_dj} is the loading (or weighting/importance) of the d^{th} common factor for environment j , $f_{a_{di}}$ is the regression coefficient (or response) of genotype i to the d^{th} common factor, and $\delta_{a_{ij}}$ is the residual genetic effects specific to environment j . Therefore, each environment is characterised in the model by the loadings of the c common factors, plus a residual genetic effect which captures genetic variation not explained by the common factors. The performance of any genotype in environment j is a function of its response to the c common factors, plus the residual genetic effect specific for environment j . The variance of the common factors (f_a) and residual genetic effects (δ_a) are assumed to follow: $\begin{bmatrix} f_a \\ \delta_a \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{I}_{c_a} \otimes \mathbf{A} & 0 \\ 0 & \boldsymbol{\Psi}_a \otimes \mathbf{A} \end{bmatrix} \right)$, where $\mathbf{I}_{c_a}$ is an identity matrix of order c , and $\boldsymbol{\Psi}_a$ is a $p \times p$ diagonal matrix with elements ψ_{aj} , which are the residual genetic variances for the p environments. The genetic covariance matrix between the p environments is estimated using $\mathbf{G} = \mathbf{\Lambda}_a \mathbf{\Lambda}_a' + \boldsymbol{\Psi}_a$, where $\mathbf{\Lambda}_a$ is a $p \times c$ matrix of estimated loadings for the p environments, and $\boldsymbol{\Psi}_a$ is a $p \times p$ matrix with the residual genetic variances for individual environments on the diagonal. As the number of fitted common factors increases, the proportion of genetic variation captured by the residual genetic variance will decrease. The number of common factors required to adequately explain the genetic (co)variance between environments is typically selected based on a combination of goodness-of-fit criteria and the overall percentage of the average genetic variance explained across environments (Cullis et al., 2014; Smith & Cullis, 2018; Tolhurst et al., 2019). Only a small number of common factors are usually required to achieve this; for example, Smith et al. (2021) required four common factors to explain 74.2% of the genetic variance across 73 diverse environments, while the residual genetic effects explained the remaining 25.8%. In situations where the residual genetic effects are not significant, a reduced rank parameterisation of the factor analytic model can be fit, which assumes that the residual genetic effects are zero.

A substantial advantage of the factor analytic model is that it does not rely on a pre-determined EC that might be uninformative. Instead, it will estimate the ECs (i.e., common factors) such that they explain the maximum amount of the covariance between the environments. Likewise, the model is more flexible than the reaction norm, as it does not enforce smooth changes in the trait between environments (Meyer, 2009). Factor analytic models have been used to model G×E interactions in most commercially relevant crops, including wheat (Aoun et al., 2021; Sjöberg et al., 2021), barley (Ferrante et al., 2021), and maize (Dias et al., 2018; Fritsche-Neto et al., 2010), as well as trees (Cullis et al., 2014; Hardner et al., 2011). More recently, methods for extracting breeding values for the overall performance and robustness of genotypes across environments have been derived using the regressions of individual genotypes on the common factors (Smith & Cullis, 2018). However, a disadvantage is that we rarely know what each common factor represents, and the model cannot provide predictions into future (or untested) environments. This has been recently addressed through the incorporation of environmental data into factor analytic models (Tolhurst et al., 2022), which enables the prediction of variety performance in untested environments.

Only three examples of the factor analytic model applied to G×E in livestock were found in the literature. In the first, significant G×E interactions in growth traits were found in sheep across eight different locations (Li et al., 2019). The second example detected significant G×E interaction in rainbow trout across four different environments using a single common factor (Sae-Lim et al., 2014). The common factor explained on average 49.4% of the additive genetic variance across the environments and only captured variation due to scale-type G×E interaction (because the loadings for the factor were all positive). They found that a ‘day*degree’ variable was highly correlated with the common factor, implying that it explained most of the scale-type G×E interactions in the data. Finally, Tyrisev  et al. (2011) used factor analytic models to parameterise the genetic covariance between 25 different countries for protein yield, which

greatly reduced the computing time without affecting the prediction of breeding values. However, no studies of livestock have applied the tools outlined by Smith & Cullis (2018) to calculate breeding values of overall performance and robustness from these models, while their application in plants is also relatively new. Therefore, there is great potential to further explore the use of factor analytic models in G×E studies of livestock and plants.

2.4 Robustness

There is a significant body of research that discusses the definitions of robustness. Related concepts have been described for livestock that may be closely associated with robustness, such as, such as resilience, tolerance, macro-environmental sensitivity, phenotypic plasticity and stability, to name just a few (Doeschl-Wilson et al., 2021; Friggens et al., 2017; Iung et al., 2020). This can be confusing to interpret, as the terms “robustness”, “macro-environmental sensitivity” and “resilience” have all been used interchangeably to describe the slope of a reaction norm in livestock (Madsen et al., 2018; Mulder, 2016; Rauw & Gomez-Raya, 2015). For the most part, studies have attempted to define differences between the concepts descriptively. A recent example in livestock can be found in Berghof et al., (2019), which defined robustness as “Animals that combine high production potential with resilience to external stressors, allowing for unproblematic expression of high production potential in a wide variety of environmental conditions”. In contrast, macro-environmental sensitivity was defined as “the property of the organisms of a genotype to develop systematically different phenotypes in different environments = norm of reaction”. Rather than focusing on broad definitions, which can be vague and loosely interpreted, it could be more informative to use a model-based definition that is grounded in what genetic variation is being captured. In this thesis, a robust genotype is defined as having genetic merit that is less sensitive to environmental variation. That is, if a sire with robust breeding value for body weight was to be used across many different environments, the performance of its progeny for body weight would be more stable

across environments compared to a less robust sire. This definition is consistent with the variation that is captured in $G \times E$ interaction models and provides a clear link to the practical consequence of robustness. In plant breeding, the term ‘stability’ is used to mean the same thing (Kraakman et al., 2004; Sandhu et al., 2021; Smith & Cullis, 2018). This thesis will use the term ‘robustness’ instead of ‘stability’, except in chapters that are for publication in the plant breeding literature.

Although genotype by environment interactions have been studied for some time (Woltereck, 1913), the extent of genetic variation in robustness and its potential for selection in livestock breeding programs is still relatively unexplored. To our knowledge, robustness is not directly included in any livestock breeding program except for heat tolerance in Australian dairy cattle (Nguyen et al., 2017). It is clear more research is required to understand the methods of estimation for robustness and its potential application to improve livestock populations. This thesis examines $G \times E$ interactions and genetic variation in robustness utilising populations of both livestock (sheep) and plants (barley). We examine more deeply the properties of breeding values for robustness estimated using reaction norms and contrast these with factor analytic models used in plant analyses. Furthermore, we explore the predictive ability of breeding values for robustness in livestock.

Chapter 3. Genomic analysis of the slope of the reaction norm for body weight in Australian sheep

Dominic L. Waters^{1*}, Sam A. Clark¹, Nasir Moghaddar¹, and Julius H. van der Werf¹

¹School of Environmental & Rural Science, University of New England, Armidale, NSW, 2351, Australia

*Corresponding author

E-mail addresses:

DLW: dwaters9@myune.edu.au

SAC: sam.clark@une.edu.au

NM: nmoghad4@une.edu.au

JHVDW: jvanderw@une.edu.au

Published in Genetics Selection Evolution, June 2022

3.1 Abstract

Background

Selection of livestock based on their robustness or sensitivity to environmental variation could help improve the efficiency of production systems, particularly in the light of climate change. Genetic variation in robustness arises from genotype-by-environment ($G \times E$) interactions, with genotypes performing differently when animals are raised in contrasted environments. Understanding the nature of this genetic variation is essential to implement strategies to improve robustness. In this study, our aim was to explore the genetics of robustness in Australian sheep to different growth environments using linear reaction norm models (RNM), with post-weaning weight records of 22,513 lambs and 60k single nucleotide polymorphisms (SNPs). The use of scale-corrected genomic estimated breeding values (GEBV) for the slope to account for scale-type $G \times E$ interactions was also investigated.

Results

Additive genetic variance was observed for the slope of the RNM, with genetic correlations between low- and high-growth environments indicating substantial re-ranking of genotypes (0.44-0.49). The genetic variance increased from low- to high-growth environments. The heritability of post-weaning body weight ranged from 0.28 to 0.39. The genetic correlation between intercept and slope of the reaction norm for post-weaning body weight was low to moderate when based on the estimated (co)variance components but was much higher when based on back-solved SNP effects. An initial analysis suggested that a region on chromosome 11 affected both the intercept and the slope, but when the GEBV for the slope were conditioned on the GEBV for the intercept to remove the effect of scale-type $G \times E$ interactions on SNP effects for robustness, a single genomic region on chromosome 7 was found to be associated

with robustness. This region included genes previously associated with growth traits and disease susceptibility in livestock.

Conclusions

This study shows a significant genetic variation in the slope of RNM that could be used for selecting for increased robustness of sheep. Both scale-type and rank-type $G \times E$ interactions contributed to variation in the slope. The correction for scale effects of GEBV for the slope should be considered when analysing robustness using RNM. Overall, robustness appears to be a highly polygenic trait.

3.2 Background

Genotype-by-environment ($G \times E$) interactions occur when the expression of the genes carried by an individual changes depending on the environment in which it is raised. This means that the best animal for one environment might not necessarily be the best one in another environment. One approach to this problem is to design breeding programs for specific geographical environments (Mulder & Bijma, 2005). However, within the same geographical location, environments often vary considerably from year to year, such as during periods of drought (Nicholls et al., 1997). An alternative is to select animals that rank highly across wide-ranging environmental conditions. These animals are said to be robust, as their breeding value is less sensitive to the environment.

The robustness of an animal cannot be directly measured, since it is typically raised in a single location at a given time point. However, robustness can be inferred for a genotype using a reaction norm model (RNM) that includes genetic relationships between relatives that have been raised in different environments. In a RNM, the estimated breeding value (EBV) of each animal is modelled as a function of a continuous environmental descriptor using random

regression. A linear function is typically used as it is easier to interpret than higher order polynomials (Kolmodin et al., 2002). This results in two EBV for each individual, i.e., an intercept, which captures the average performance across environments, and a slope, which captures the amount of change in EBV across the environmental descriptor. The EBV for the slope is directly treated as an EBV for robustness. Many previous studies have used RNM to model sensitivity to environmental changes (Kolmodin et al., 2002; Madsen et al., 2018). For example, Hollema et al. (2018) used a linear RNM to model variation in the sensitivity of Australian sheep to parasite burden using pedigree relationships. The average worm-burden of a contemporary group was used as the environmental descriptor, and a flatter slope was interpreted as a more robust genotype.

More recently, RNM have been extended to include genomic information in the form of single nucleotide polymorphisms (SNPs) (Silva et al., 2014). The use of genomic data increases the power for estimating robustness, since it captures the genetic linkage between distantly-related genotypes in different environments (Hayes et al., 2016). Compared to regular genomic best linear unbiased prediction (GBLUP) models that ignore $G \times E$ interactions, RNM models have shown increased predictive ability (Mota et al., 2020; Oliveira et al., 2018). When a reference population is well-distributed across environments, selection based on genomic estimated breeding values (GEBV) greatly improves the response to selection for robustness (Mulder, 2016), as young sires can be selected earlier without the need for large progeny tests across multiple environments. In addition, genomic RNM can be used for genome-wide association (GWA) studies by back-solving GEBV to SNP effects. This can shed light on the molecular basis underlying robustness and improve the accuracy of genomic predictions by using models which incorporate the molecular information as priors (MacLeod et al., 2016).

The correlation between the genetic variance in the intercept and that in the slope has often been limited to a discussion as a correlation between overall performance and sensitivity (Kolmodin et al., 2003). Beyond this, a correlation implies that (1) scale changes in genetic variance occur across environments, and (2) information from the intercept is used to estimate the EBV for the slope (and vice versa). Both of these observations could influence the EBV for the slope but are unrelated to the re-ranking of genotypes. Therefore, it might be important to ensure a level of independence between the EBV for the slope and the EBV for the intercept in the analysis of robustness.

The objectives of this study were to: (1) estimate the genetic variation of robustness in Australian sheep between different growth environments using linear reaction norms models fitted by random regression; and (2) compare the use of direct and scale-corrected GEBV for the slope in a genome-wide association study, using post-weaning weight (PWWT) as the response variable.

3.3 Methods

3.3.1 Data structure

In this study, data from 34,584 sheep recorded in the Australian Sheep CRC Information Nucleus Flock (INF) and the Meat and Livestock Australia Resource Flock (RF) were used. The INF animals were located at eight sites across southern Australia between 2007 and 2011. Each year, approximately 100 sires from the two main sites (Armidale NSW and Katanning WA) were mated via artificial insemination, each one with 900 ewes. For the six remaining sites, 450 ewes were mated annually, with at least 50% of the Kirby/Katanning sires. Then from 2012 onwards, the program was changed: it used the RF and only the Armidale and Katanning sites were kept under a similar mating structure, but with about 150 sires mated at

both sites each year. The management practices at each site reflected the typical practices used for raising lambs on pasture in each of the respective geographical areas. Supplemental feed was occasionally provided in years with periods of drought to help maintain the body weight of animals.

The population consisted of multiple breeds. Merino, Maternal (such as Border Leicester) or Terminal (such as Poll Dorset) sires were mated to either Merino or first cross Maternal-Merino dams. Only the Merino breed had pure-bred representatives, while the remaining breeds were represented as crosses. Overall, 39 different breed groups were available, referred to as genetic groups hereafter. A summary of the breed composition of the data is Appendix A, Table A3.1. For more information on the INF and RF flocks, see van der Werf et al. (2010).

Phenotypic records for post-weaning growth rate (PWGR) and post-weaning weight (PWWT) were used. First, PWGR was used to estimate the environmental descriptor used in the RNM. The PWGR of each lamb was calculated as the difference between weaning and post-weaning weights, recorded at approximately 96 (64-120) and 238 (120-329) days of age, respectively, which was divided by the number of days between measurements and expressed in g/day. PWWT was then used as the dependent variable for the RNM analysis. The distribution of PWGR and PWWT as well as the other descriptive statistics after filtering are available in Appendix A, Table A3.2.

3.3.2 Preparation of the reaction norm data

Since explicit environmental information was not available for each of the sites, the best linear unbiased estimation (BLUE) of the contemporary group effect for PWGR was used to estimate the environmental descriptor. Contemporary groups (CG) were formed as a cohort of site×birth year×management group. The management group consisted of animals which were

subjected to the same management decisions within each site×year, i.e., they were raised in the same paddocks and phenotyped at the same time. PWGR was chosen as the environmental descriptor because rate of growth has an intuitive relationship with environmental quality, i.e., CG that grow more quickly after weaning, are more likely to be in a better environment than groups that grow more slowly. When implemented in a reaction norm model, the slope of a genotype is interpreted as the change in GEBV per change in growth environment (g/day). Other studies have also demonstrated the value in using alternative traits for the environmental gradient (EG), rather than the trait used as the reaction norm response variable (i.e., PWWT) (Guy et al., 2019).

Lambs that had a growth period shorter than 40 days, a PWGR more than 4 standard deviations (SD) from the mean PWGR or that were born as quintuplets or reared as quadruplets were not included in the analysis. Only CG with a minimum of 15 animals were considered. After filtering, 33,773 individuals from 249 CG were available to estimate the CG effects. The model for obtaining the BLUE of CG effects was a regular pedigree-based animal model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a} + \mathbf{Q}\mathbf{g} + \mathbf{e} \quad (1)$$

where $\mathbf{y}$ is the vector of PWGR records, $\mathbf{X}$ is an incidence matrix for the fixed effects $\mathbf{b}$, $\mathbf{Z}_1$ is an incidence matrix relating the records to additive genetic effects ($\mathbf{a}$), $\mathbf{Q}$ is a matrix of the proportion of each animal's genome originating from the 39 genetic groups (Westell et al., 1988), $\mathbf{g}$ is the vector of the random genetic group effects, and $\mathbf{e}$ is the vector of the residual effects. Fixed effects included sex, birth-type and rear-type interaction, age at post-weaning (linear and quadratic) and CG.

Only the CG effects for genotyped animals (22,634) were extracted. Each animal was assigned the BLUE of their CG as their environmental descriptor, which was then standardised to have

a mean of 0 and variance of 1, forming the EG used in the RNM. This subset of genotyped animals was re-filtered to ensure that the animals had a PWWT within 4 SD from the mean PWWT, originated from CG of at least 15 other animals, or had at least one half-sib. After filtering, 22,513 animals from 1582 sires and 11,576 dams across 206 CG were available for RNM analysis.

The CG effects ranged from -2.15 SD (-72.4 g/day) to 2.06 SD (69.1 g/day) respectively, with an overall range of 141.5 g/day. On average, sires had progeny in environments with PWWT differing by 59 g/day. A small number of the sires (6%) had progeny in only one environment, but these were still included since their genomic data can link them to other animals across environments. The distribution of records and mean PWWT along the EG is plotted in Appendix A, Figure A3.1.

3.3.3 Genotype data

Not all animals were genotyped due budget constraints. Genotyped animals were selected randomly from each sire to minimise selection bias. The imputed genomic data for each individual consisted of 60,400 SNPs. Originally, animals were genotyped with a 12k, or one of two 15k or one of two 50k SNP panels. Approximately 93% of the animals were genotyped with one of the 50k SNP panels. Within each panel, SNPs were removed if they were located on the X or Y chromosome, deviated greatly from the Hardy-Weinberg equilibrium ($P < 10^{-15}$), had call rates lower than 90% and Gen-Cal scores lower than 0.6. When the correlation between the genotype sample of an animal and that of another animal was higher than 0.98, they were removed, as this indicated an error in matching ID. Following quality control, the Beagle 5.1 software (Browning et al., 2018) was used to impute all SNP genotypes to the combined set of 60,400 SNPs. The number of reference animals used to impute the SNP genotypes for each chip is in Appendix A, Table A3.3. After imputation, the combined set was

filtered to remove SNPs with a minor allele frequency lower than 0.01. Finally, 60,347 SNPs were available for analysis.

3.3.4 Reaction norm analysis

Two RNM were used to estimate GEBV. The first RNM (RNM-HOM) assumed that the residual variance was homogenous along the EG, and is described by:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a}_1 + \mathbf{Z}_2\mathbf{a}_2 + \mathbf{Q}\mathbf{g} + \mathbf{e}, \quad (2)$$

where $\mathbf{y}$ is the vector of PWWT records, $\mathbf{X}$ is an incidence matrix linking records to the fixed effects $\mathbf{b}$, and $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are incidence matrices linking records to the breeding values for the intercept ($\mathbf{a}_1$) and the slope ($\mathbf{a}_2$), $\mathbf{Q}$ is a matrix of the proportion of each animal's genome originating from the 39 genetic groups, $\mathbf{g}$ is the vector of the random genetic group effects, and $\mathbf{e}$ is the vector of residual effects. Fixed effects included sex, birth type, rear type, age at post-weaning and CG. The $\mathbf{Q}$ matrix and CG were formed as previously described for the BLUE model.

The genetic variance of $\mathbf{a}_1$ and $\mathbf{a}_2$ was modelled according to $\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix} \sim N(\mathbf{0}, \mathbf{G} \otimes \mathbf{K})$ where $\mathbf{K} = \begin{bmatrix} \sigma_{a_1}^2 & \sigma_{a_2 a_1} \\ \sigma_{a_1 a_2} & \sigma_{a_2}^2 \end{bmatrix}$ and $\mathbf{G}$ is a relationship matrix based on genomic data, constructed following the first method described by VanRaden (2008). Genetic groups were modelled as for Eq. (1). The variance of the residual $\mathbf{e}$ was modelled according to $[\mathbf{e}] \sim N(\mathbf{0}, \mathbf{I}\sigma_e^2)$, where σ_e^2 is the residual variance which was assumed to be homogenous across environments.

The second RNM (RNM-HET) was the same as RNM-HOM, except that residual variance was modelled as a function of the EG, as described in (Ni et al., 2019). An intercept ($\mathbf{e}_1$) and slope

($\mathbf{e}_2$) residual coefficient was used, such that $\begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} \sim N(\mathbf{0}, \mathbf{I} \otimes \mathbf{E})$, where $\mathbf{E} = \begin{bmatrix} \sigma_{e_1}^2 & \sigma_{e_2 e_1} \\ \sigma_{e_1 e_2} & \sigma_{e_2}^2 \end{bmatrix}$.

The models were implemented using genomic residual maximum likelihood (GREML) with the MTG2 software (Lee & van der Werf, 2016). The genetic (co)variance matrix for breeding values between environments along the EG were estimated as: $\hat{\mathbf{V}} = \mathbf{\Lambda} \mathbf{K} \mathbf{\Lambda}'$, where $\mathbf{K}$ is the additive genetic (co)variance matrix for the intercept and the slope, and $\mathbf{\Lambda}$ is a 100×2 matrix containing a vector of 1s for the intercept and a vector of 100 standardised environmental values ranging from the minimum to the maximum value of the EG. The genetic correlation matrix between environments was computed from $\hat{\mathbf{V}}$.

For RNM-HET, the residual variance for environment at level i was calculated as the element i, i in the matrix $\hat{\mathbf{R}}$, which was estimated as: $\hat{\mathbf{R}} = \mathbf{\Lambda} \mathbf{E} \mathbf{\Lambda}'$, where $\mathbf{E}$ is the (co)variance matrix for the residual intercept and slope regression coefficients, and $\mathbf{\Lambda}$ is the same as above. The heritability of PWWT in environment i was estimated as: $h_i^2 = \frac{V_{a_i}}{V_{a_i} + V_{e_i}}$, where V_{a_i} is the additive genetic variance estimated for environment i and V_{e_i} is the residual variance in environment i . Standard errors for the variance components and heritabilities were estimated using a Taylor series expansion (Fischer et al., 2004).

To observe how parameters changed along the EG, the genetic variance, correlation and heritabilities from the two models were compared between three distinct environments: -1.5 SD, 0 SD and 1.5 SD, referring to low-, average- and high-growth environments, respectively. Evaluation of changes in parameters was limited to this range of EG values, as estimates outside this range may be biased since they are based on relatively small amounts of data (Meyer, 2005). To further validate the reaction norm variance parameters, the data were split into three sections based on the EG: low (< -1.13 SD), average (> 0.59 SD and < 0.59 SD) and high (> 1.13 SD), such that the mean EG of each section was -1.55, 0.06 and 1.52 SD for low, average

and high sections, respectively. In total, 3157, 9375 and 3384 animals were included in the low, average and high sections, respectively. A multi-trait model (MTM), which considered performance in the three sections as different traits but allowed them to be genetically correlated, was then fitted. This provided a robust estimate of the parameters for the three distinct environments, which could be used as a benchmark for the reaction norms, as done in (Calus & Veerkamp, 2003).

3.3.5 Genome-wide association (GWA) analysis

To obtain scale-corrected GEBV for the slope ($\mathbf{a}_2^*$) from RNM-HOM and RNM-HET for GWA analysis, the following genetic regression was used:

$$\mathbf{a}_2^* = \mathbf{a}_2 - \frac{\sigma_{a_1 a_2}}{\sigma_{a_1}^2} \mathbf{a}_1, \quad (3)$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ are the GEBV for the intercept and the slope, respectively, $\sigma_{a_1}^2$ is the variance in intercept and $\sigma_{a_1 a_2}$ is the covariance between intercept and slope, estimated from the respective RNM models. The aim of this formula was to remove the variation in GEBV for the slope that originated from scale effects. To distinguish the different GEBV for the slope, the terms direct slope ($\mathbf{a}_2$) and scale-corrected slope ($\mathbf{a}_2^*$) are used.

The GWA analyses were performed by separately back-solving the GEBV for the intercept ($\mathbf{a}_1$), the direct slope ($\mathbf{a}_2$) and scale-corrected slope ($\mathbf{a}_2^*$) to SNP effects for each model, following (Strandén & Garrick, 2009):

$$\hat{\mathbf{u}} = \frac{\mathbf{W} \mathbf{G}^{-1} \mathbf{a}}{d}, \quad (4)$$

where $\hat{\mathbf{u}}$ is the vector of SNP effects, d is a scalar calculated as $2 \sum_{i=1}^{60347} p(1-p)$, where p is the allele frequency for SNP i , $\mathbf{W}$ is the SNP matrix corrected for allele frequency differences ($\mathbf{W} = \mathbf{M} - 2 * (p - 0.5)$, $\mathbf{M}$ is the SNP matrix coded such that 0 represents the heterozygous genotype, and -1 and 1 represents the homozygous genotypes), $\mathbf{G}^{-1}$ is the inverse of the genomic relationship matrix (VanRaden, 2008), and $\mathbf{a}$ is the vector of GEBV (either $\mathbf{a}_1$, $\mathbf{a}_2$, or $\mathbf{a}_2^*$).

An approximate p -value of each SNP effect was estimated based on a t-distribution, calculated as the probability of the 95th percentile of the GEBV distribution (Gondro, 2015). A Bonferroni correction was used to avoid false positives. A SNP was considered significant when its p -value was lower than $0.05/60,347$, giving a threshold of $-\log_{10}(p) > 6.08$. Genes in the genomic regions that were significantly associated with the scale-corrected slope were detected based on the *Ovis aries* v3.1 assembly using the NCBI genome data viewer (Agarwala et al., 2018).

3.4 Results

3.4.1 Reaction norm models

Additive genetic variation was observed for the slope, indicating the presence of G×E interactions and genetic variation in robustness for PWWT (Table 3.1). The genetic correlation between the intercept and the slope, obtained from the additive genetic (co)variance components, was much lower in RNM-HET (0.18) than in RNM-HOM (0.52). Overall, RNM-HET provided a better fit based on the log-likelihood, which indicated that the residual variance was heterogenous along the EG.

Table 3. 1 Variance of the additive genetic intercept (σ_{a1}^2) and slope (σ_{a2}^2), the residual intercept (σ_{e1}^2), slope (σ_{e2}^2) and covariance ($\sigma_{e1}\sigma_{e2}$), as well as genetic group (g) variance. The correlation between additive genetic intercept and slope (r_{a1a2}) and the log-likelihood (LKH) are also reported

Model ^a	σ_{a1}^2	σ_{a2}^2	r_{a1a2}	σ_{e1}^2	σ_{e2}^2	σ_{e1e2}	g	LKH
RNM-	7.40	1.28	0.52	17.52	-	-	11.45	-46938
HOM	(0.31)	(0.16)	(0.04)	(0.26)			(3.85)	
RNM-	6.98	1.19	0.18	18.17	-0.32	1.55	10.88	-46870*
HET	(0.31)	(0.17)	(0.06)	(0.31)	(0.23)	(0.14)	(3.69)	

^aRNM-HOM: linear reaction norm model with homogenous residual variance; RNM-HET: linear reaction norm model

with heterogenous residual variance. *Based on a log-likelihood ratio test using a chi-square distribution, RNM-HET provided a better fit ($p = 1.7 \times 10^{15}$)

The genetic variance increased by 177% from low- to high-growth environments in RNM-HOM (Figure 3.1a) but increased only by 39% across the same environments in RNM-HET, although the greatest increase in RNM-HET was between average- and high-growth environments (61%). This still showed substantially lower scale-type G×E interactions and indicated considerable bias in the genetic parameters of RNM-HOM. The estimates of heritability in RNM-HOM followed the same pattern as the genetic variance, increasing by 92% from low- to high-growth environments (Figure 3.1b). The heritability in RNM-HET was higher in low-growth environments (0.39) than in average- and high-growth environments (0.28-0.33). The residual variance increased by 74% from low- to high-growth environments in RNM-HET (see Appendix A, Figure A3.2).

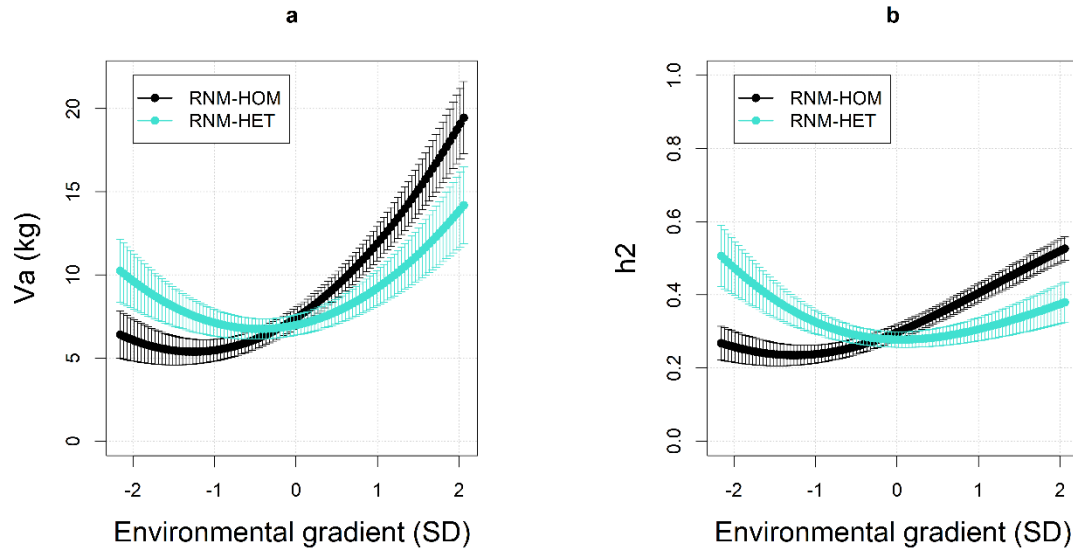


Figure 3. 1 Estimates of (a) genetic variance (V_a) and (b) heritability (h^2) with 95% confidence intervals along the environmental gradient

Genetic correlations between low- and high-growth environments were 0.49 and 0.44 in RNM-HOM and RNM-HET, respectively (Figure 3.2), which indicated a high degree of rank-type $G \times E$ interactions along the EG and variation in robustness. Overall, both models produced similar estimates of rank-type $G \times E$ interactions.

The parameter estimates of RNM-HET were largely consistent with those of the MTM (see Appendix A, Table A3.4-6). However, the genetic correlations obtained with the MTM were slightly lower between similar environments (Low vs Average, Average vs High) and higher between the less similar environments (Low vs High) than with RNM-HET. The parameter estimates of RNM-HOM were substantially different to those of the MTM, which suggests that RNM-HOM was biased.

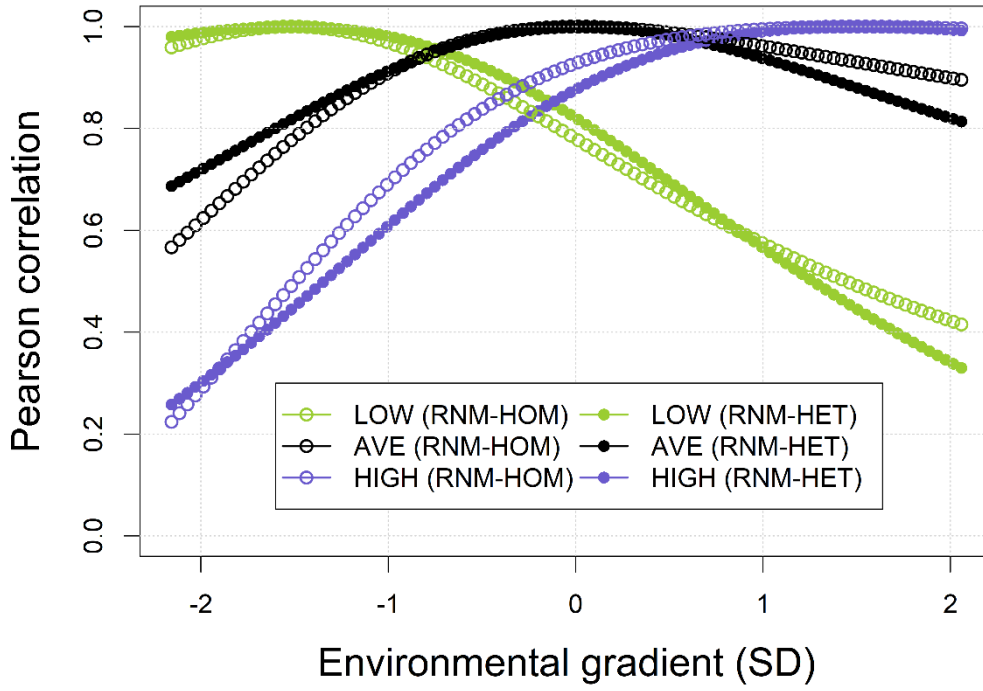


Figure 3. 2 Estimates of genetic correlations between low- (– 1.5 SD, green), average- (0 SD, black), and high- (1.5 SD, purple) growth environments from the two models.

3.4.2 GWA analysis and scale correction

The correlation between back-solved SNP effects for the intercept and the direct slope in RNM-HOM was high (0.85), indicating a large amount of scale-type $G \times E$ interactions (Figure 3.3a). The correlation was lower in RNM-HET (0.31), but still suggested scale effects (Figure 3.3c). These correlations were considerably higher than the genetic correlations based on the estimated (co)variance components for the intercept and the slope in both models in Table 1 (RNM-HOM 0.85 vs 0.52, RNM-HET 0.31 vs 0.18).

To estimate the scale-corrected SNP effects following Eq. (3), the correction formulas for the scale-corrected GEBV of RNM-HOM and RNM-HET were: $\mathbf{a}_2^* = \mathbf{a}_2 - 0.21\mathbf{a}_1$ and $\mathbf{a}_2^* = \mathbf{a}_2 - 0.075\mathbf{a}_1$, respectively. Due to the higher genetic correlation between the intercept and the slope in RNM-HOM, the GEBV and subsequent SNP effects for the slope were more strongly

corrected than in RNM-HET, because they contained more scaling effects. Once back-solved, the scale-corrected SNP effects for the slope were far less correlated with the intercept for both RNM-HOM (0.11) and RNM-HET (-0.06) (Figure 3.3b and d). Therefore, the correction appeared to remove the effect of scale-type $G \times E$ interactions on the SNP effects for the slope.

A region that was significantly associated with the scale-corrected slope (described later) is highlighted in red in all the plots in Figure 3.3. Before the correction, several SNPs in this region were associated with the intercept, especially in RNM-HOM. The correction uncovered this region in RNM-HOM by lifting away the associations of these SNPs. This also occurred in RNM-HET, but to a lesser extent.

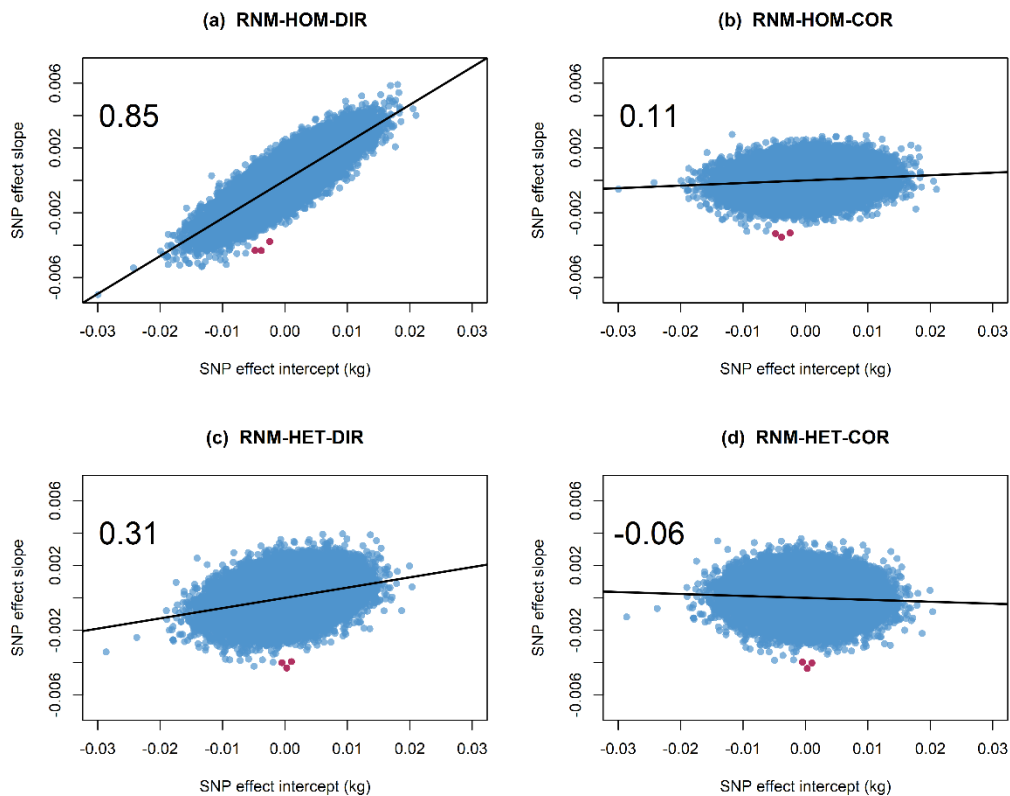


Figure 3. 3 Scatterplots and correlations between SNP effects for the direct slope (a and c) and scale-corrected slope (b and d) against the intercept. Significant SNPs on chromosome 7 are highlighted in red

The effect of the scale correction was also observed in the genomic reaction norms of genotypes before and after correction (Figure 3.4). Genotypes with a large intercept effect had a greater correction in the slope. The sign of the slope also changed after correction for some genotypes.

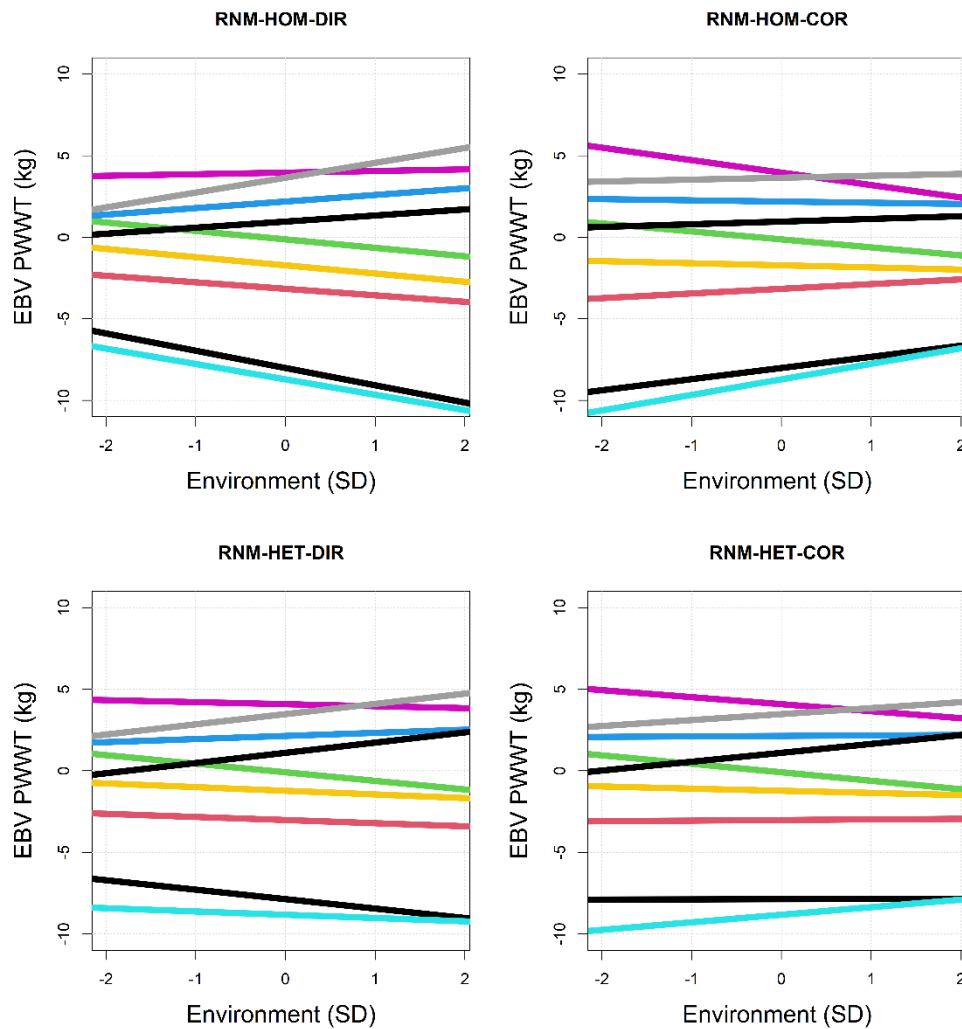


Figure 3. 4 Reaction norms before and after scale-correction for nine individuals representing the nine largest half-sib families with the largest distribution along the environmental gradient. Individuals are distinguished by colour

In RNM-HOM, Manhattan plots for the direct and scale-corrected slope detected different signals (Figure 3.5a-b). The direct slope yielded a signal on chromosome 11, which was also associated with the intercept (see Appendix A, Figure A3.3), but it disappeared when using the scale-corrected slope, and a new signal was detected on chromosome 7.

The same region on chromosome 7 was detected using the direct slope from RNM-HET, although it was not significant. However, this region was significant when using the scale-corrected slope in RNM-HET. In spite of significant differences between the Manhattan plots of the two models using the direct slope, the Manhattan plots for scale-corrected slope from both models were very similar, and the correlation between their SNP effects was very high (0.99, (see Appendix A, Figure A3.4). Overall, the GWA analysis suggests that robustness is a highly polygenic trait that is regulated by many genes with small effects.

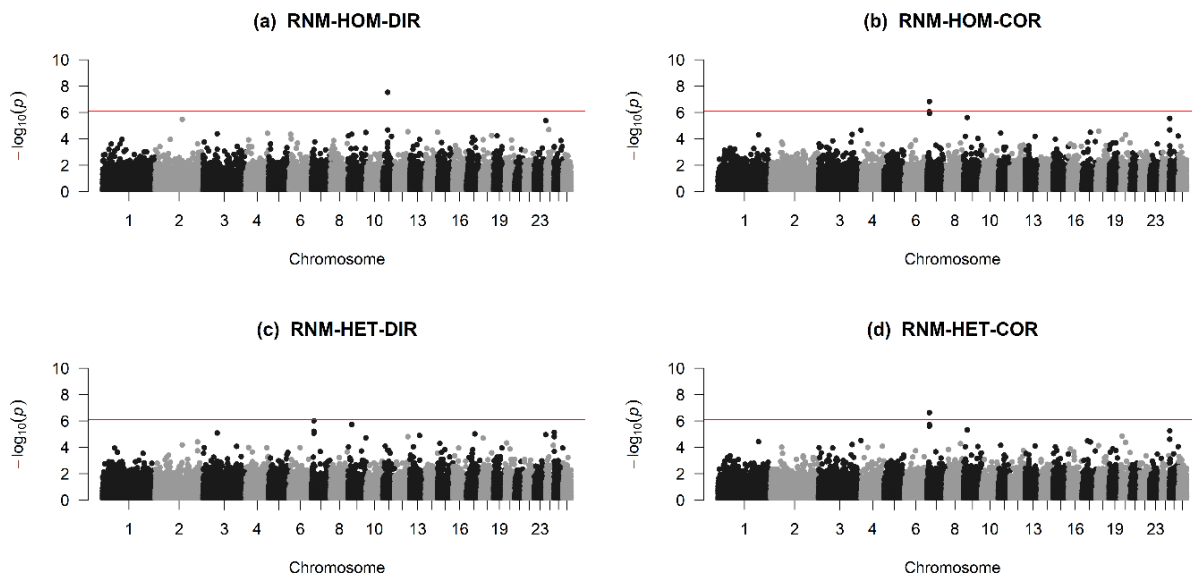


Figure 3. 5 Genome-wide SNP associations for the slope using direct (a and c) and scale-corrected (b and d) GEBV for the slope, for both RNM-HOM and RNM-HET

The three most significant SNPs associated with robustness were within an 881-bp region on chromosome 7 (Table 3.2). The allele substitution effect for these SNPs indicated that the animals with genotypes carrying the minor allele were approximately 4 g heavier at post-weaning for each one SD decrease in EG, compared to genotypes carrying the major allele. Only one of these SNPs was significant based on the Bonferroni-corrected p-value, which indicates a weak association. The three SNPs were located within the *SLC24A1* gene, which

encodes a sodium/potassium/calcium exchanger. Eight other genes were located within a 500-kb region surrounding the most significant SNP (Table 3.3).

Table 3. 2 SNPs associated with level of robustness

Chromosome	Base pair	MAF	$-\log_{10}(p)$	ASE (g / SD EG)
7	12,312,079	0.49	5.73	-4.0
7	12,312,384	0.39	5.61	-4.0
7	12,312,960	0.49	6.63	-4.4

MAF: Minor allele frequency

p: Bonferroni-corrected p-value

ASE: Allele substitution effect, expressed in g per unit SD of the EG

Table 3. 3 List of known genes in a 500-kb window surrounding the most significant SNP

Gene symbol	NCBI gene ID	Distance from the most significant SNP (kb)	Gene description
<i>SLC24A1</i>	101121857	0	sodium/potassium/calcium exchanger
<i>INTS14</i>	101121097	30.0	integrator complex subunit 14
<i>HACD3</i>	101120835	59.5	3-hydroxyacyl-CoA dehydratase 3
<i>DPP8</i>	101119728	106.5	dipeptidyl peptidase 8
<i>IGDCC4</i>	101111715	189.1	immunoglobulin superfamily DCC subclass member 4
<i>IGDCC3</i>	101111458	222.3	immunoglobulin superfamily DCC subclass member 3
<i>DENND4A</i>	101122440	138.6	C-myc promoter-binding protein
<i>RAB11A</i>	780464	212.1	ras-related protein Rab-11A
<i>MEGF11</i>	101123125	240.6	multiple epidermal growth factor-like domains protein 11

3.5 Discussion

3.5.1 Reaction norm models

Additive genetic variation was observed for the slope of the RNM, which indicated the presence of G×E interactions for PWWT in different growth environments. Both scale- and rank-type G×E interactions appeared to contribute to the variation in the slope, as evidenced by the differences in genetic variance along the EG and the much lower than 1 genetic

correlations between environments. Other studies have identified significant G×E interactions for body weight of sheep using RNM (Dominik et al., 2013; Hollema et al., 2018; Santana et al., 2013), although the genetic correlations between environments found in our study are slightly lower than those in previous studies. The estimates of heritability for PWWT in the best model, i.e., RNM-HET, were comparable to those obtained in previous studies for body weight in Australian sheep (0.31 to 0.39) (Mortimer et al., 2017; Pollot & Greeff, 2004). Overall, our findings demonstrate variation in the performance of sheep between growth environments, which can be captured using RNM.

We found that fitting heterogeneous residual variance significantly improved the fit of the RNM, as reported in similar RNM studies of livestock (Li & Hermes, 2016; Madsen et al., 2018; Santana et al., 2013). When homogeneous residual variance was assumed, the genetic variance increased by approximately 177% along the EG, which greatly over-estimated the amount of scale-type G×E interaction. Scale-type G×E interaction was detected at a much lower level after fitting heterogeneous residual variance. This agrees with (Lillehammer et al., 2009), who found that not accounting for existing heterogeneous residual variance inflated the scale-type G×E interactions. In spite of the difference in scale-type G×E interactions, estimates of genetic correlation between environments were similar for the two models, RNM-HOM and RNM-HET. The low genetic correlation (0.44-0.49) suggested a large amount of re-ranking between environments, which was evident in Figure 3.4, when the reaction norms of genotypes crossed over along the EG.

The genetic correlation between the intercept and the slope is commonly interpreted as a genetic correlation between overall performance and robustness, i.e., a high correlation indicates that genotypes that perform well tend to be more sensitive (Knap, 2005; Rauw & Gomez-Raya, 2015). This interpretation is problematic, as low-performing genotypes with

negative intercepts will also have proportionally larger slopes (i.e., are equally as sensitive). Alternatively, a genetic correlation between the intercept and the slope could simply represent a scale effect, where some of the variance in GEBV along the EG is due to heterogeneous genetic variance. This variation would not involve re-ranking and therefore might not reflect robustness. Thus, we used genetic regression to account for scale, where the GEBV for the slope and subsequent SNP effects were adjusted so that they were independent of the genetic correlation with the intercept. The same approach has been used in the analysis of residual feed intake (van der Werf, 2004).

3.5.2 Genome-wide association analysis

The direct GEBV for the slope have been used in GWA studies to identify SNPs associated with robustness (Carvalho et al., 2019; Silva et al., 2014; Zhang et al., 2019). However, this approach could result in SNPs that are associated with the intercept masking the SNPs associated with re-ranking. This was demonstrated in RNM-HOM, for which the high correlation between SNP effects for the intercept and slope (0.85) masked the SNPs associated with reranking (Figure 3.3a) and identified a region on chromosome 11 that affects robustness. Adjusting the GEBV to account for the scale-type $G \times E$ interactions by applying genetic regression reduced the correlation between the SNP effects considerably (0.11) and removed the signal on chromosome 11. The GWA analysis was able to reveal SNPs on chromosome 7 that were associated with robustness, independent of the scale effects. This phenomenon was also observed in RNM-HET, although it was less affected by scale. In this case, the region on chromosome 7 was non-significant when using the direct GEBV for the slope ($-\log_{10}[p] = 5.98$) but was significant after scale-correction ($-\log_{10}[p] = 6.63$). These findings suggest that correcting for scale effects on the GEBV for the slope can greatly influence the genomic regions detected and the interpretation of GWA analyses for robustness.

Interestingly, the correlation between the SNP effects for the intercept and scale-corrected slope in both models was not exactly zero. However, the genetic regression used for the scale-corrected slope was based on the estimated genetic (co)variance components, rather than the (co)variance of the GEBV used to back-solve the SNP effects. These (co)variances are not identical, as the genetic (co)variance components estimate the variance of the true breeding values (V_a), while the variance of the GEBV is equal to $V_a \times r^2$, where r is the accuracy of the GEBV. Unless the accuracies of the GEBV are all perfect (i.e., 1), the scale-correction might not result in a correlation of 0 between the back-solved SNP effects for the intercept and the scale-corrected slope. This could explain the small deviations from 0 found in this study.

Intuitively, the correlation between SNP effects estimated from the GEBV for the intercept and the slope are expected to be similar to the genetic correlation estimated in the RNM. However, the correlation between SNP effects for the intercept and the slope was considerably higher in both models, which may be explained by the fact that random regression models do not estimate GEBV for the intercept and the slope independently. Similar to a multi-trait model, they allow information from one trait (the intercept) to inform the GEBV of the other trait (the slope) through the covariance. This sharing of information can result in GEBV for the two traits having a higher correlation than the estimated genetic correlation (van der Werf et al., 2018). It is conceivable that the RNM ‘borrowed’ a large amount of information from the intercept to estimate the GEBV for the slope, especially considering that there was more variation in the intercept than in the slope. This could explain the higher correlation of SNP effects that were back-solved from the GEBV compared to the genetic correlation estimated in the RNM. Unlike a multi-trait model, in which the two traits can be analysed independently, the intercept and slope of the reaction norm must be estimated simultaneously. Therefore, using the genetic regression which accounts for this correlation could be especially important when using random regression models to study the genomics of reaction norms.

One genomic region on chromosome 7 was significantly associated with the robustness of PWWT to different growth environments. Several genes identified in this region are associated with body weight and composition-related traits in livestock. The *SLC24A1*, *DENND4A* and *RAB11A* genes are associated with both average daily gain in Nellore cattle (Olivieri et al., 2016) and back-fat thickness in a composite cattle breed (Hay & Roberts, 2018). Similarly, *MEGF11* is associated with feed efficiency in cattle and pigs (Ding et al., 2018; Vanvanhossou et al., 2020), and body length in cattle (Vanvanhossou et al., 2020). Intriguingly, some of these genes have been implicated in disease-related traits. For example, Wang et al (Wang et al., 2020) showed that *SLC24A1* is involved in respiratory and metabolic alkalosis in chickens under heat stress, and speculated that it may have a driving role in stabilizing the acid-base balance in the blood. *SLC24A1* has also been shown to be downregulated in cows with subclinical mastitis (Cheng et al., 2021), while *MEGF11* is associated with somatic cell count in cattle (Strillacci et al., 2014). In addition, the *IGDCC4* and *IGDCC3* genes encode antibodies (Agarwala et al., 2018), which are a critical component of the immune system. All of these pathways could be important when animals are raised in challenging environments.

The identification of these genes improves our understanding of the biology of robustness. However, the value of this information for increasing the accuracy of prediction of robustness is likely negligible, since the effect of these genes appears to be very small. Implementing selection for robustness based on GBLUP should efficiently capture these polygenic effects without the need to understand the underlying biology (Clark et al., 2011).

While the results of this study demonstrate the importance of considering the impact of scale-type G×E interactions when analysing robustness, the impact of the scale-correction in affecting selection decisions and genetic progress in breeding programs is less clear. A genotype which ranks highly across environments for traits such as body weight is evidently

more valuable than a genotype which only ranks highly in specific environments. Simulations that explore selection for robustness under various scenarios of G×E interactions are needed to further understand the impact of the scale-correction, as well as the value of selecting for robustness in breeding programs.

3.6 Conclusions

This study demonstrates the existence of important G×E interactions for PWWT of sheep in different growth environments and highlights the possibility of selecting sheep based on their robustness. Correction of slope breeding values using a genetic regression is suggested to account for scale-type G×E interactions on estimates of robustness. The genetic architecture of robustness in sheep appears to be highly polygenic.

Journal-Article Format for PhD Theses at the University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidates' contribution as indicated in the *Statement of Originality*.

	Author's Name	% of contribution
Candidate	Dominic Luke Waters	80
Other Authors	Sam Clark	9
	Nasir Moghaddar	2
	Julius van der Werf	9

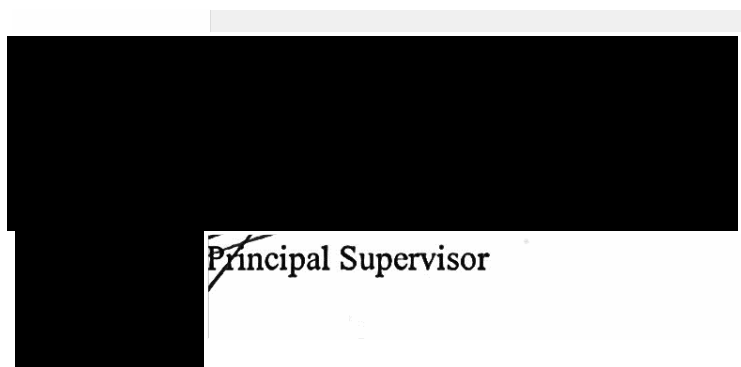
Name/title of Principal Supervisor: Professor Julius van der Werf



Candidate

30/06/2023

Date



Principal Supervisor

30/06/2023

Date

Journal-Article Format for PhD Theses at the University of New England

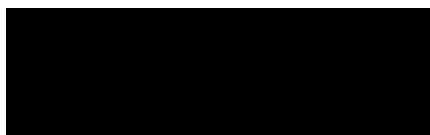
STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of Work	Page number's
Text	All
Figures	All
Tables	All

Name of Candidate: Dominic Luke Waters

Name/title of Principal Supervisor: Professor Julius van der Werf



Candidate

30/06/2023

Date



Principal Supervisor

30/06/2023

Date

Chapter 4. Partitioning the forms of genotype by environment interaction in the reaction norm analysis of stability

Dominic L. Waters^{1*}, Julius H.J. van der Werf¹, Hannah Robinson^{2,3}, Lee T. Hickey³, Sam A. Clark¹

¹School of Environmental & Rural Science, University of New England, Armidale, NSW, 2351, Australia

²InterGrain Pty Ltd, Perth, Western Australia, Australia

³ Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Brisbane, Queensland, Australia

*Corresponding author.

Email addresses:

DLW: dwaters9@myune.edu.au

JHVDW: jvanderw@une.edu.au

HR: hrobinson@intergrain.com

LTH: l.hickey@uq.edu.au

SAC: sam.clark@une.edu.au

Published in Theoretical and Applied Genetics, April 2023

4.1 Abstract

The slope of regression in a reaction norm model, where the performance of a genotype is regressed over an environmental covariable, is often used as a measure of stability of genotype performance. This method could be developed further by partitioning variation in the slope of regression into the two sources of genotype-by-environment interaction ($G \times E$) which cause it: scale-type $G \times E$ (heterogeneity of variance) and rank-type $G \times E$ (heterogeneity of correlation). Because the two types of $G \times E$ have very different properties, separating their effect would enable a clearer understanding of stability. The aim of this paper was to demonstrate two methods which seek to achieve this in reaction norm models. Reaction norm models were fit to yield data from a multi-environment trial in Barley (*Hordeum vulgare*), with the adjusted mean yield from each environment used as the environmental covariable. Stability estimated from factor analytic models, which can disentangle the two types of $G \times E$ and estimate stability based on rank-type $G \times E$, was used for comparison. Adjusting the reaction norm slope to account for scale-type $G \times E$ using a genetic regression more than tripled the correlation with factor analytic estimates of stability (0.24-0.26 to 0.80-0.85), indicating it removed variation in the reaction norm slope that originated from scale-type $G \times E$. A standardisation procedure had a more modest increase (0.55-0.59) but could be useful when curvilinear reaction norms are required. Analyses which use reaction norms to explore the stability of genotypes could gain additional insight into the mechanisms of stability by applying the methods outlined in this study.

4.2 Introduction

Genotype-by-environment interactions ($G \times E$) occur when the performance of a genotype is dependent on the environment it exists in. The performance of some genotypes can be more sensitive to environmental factors, while others show greater stability across a range of

environments. Genotypes that combine high overall performance with stability are appealing for their greater marketing potential and utility in mitigating the effects of climate change on crop production (Powell et al., 2012). Therefore, there is a need to understand the models available to estimate genetic variation and select parent genotypes for stability in plant breeding programs.

Reaction norm (RN) models have been a popular choice for modelling $G \times E$ and the stability of genotypes across environments (Kraakman et al., 2004; Kusmec et al., 2017; Lacaze et al., 2009; Li et al., 2019; Ly et al., 2018; Sjöberg et al., 2020). In RN models, the performance of a genotype is regressed across an environmental covariable (EC), which describes the quality of the environment as a continuous value (Woltereck, 1913). If the EC is centred with a mean of zero, the intercept represents performance in the average EC, while the slope of the regression captures the stability of the genotype. The regression can be either fixed or random; random regression is more common, as additive relationships can then be accounted for via pedigree or genomic data, which enables the estimation of additive genetic breeding values and genomic predictions of stability. The current paper focuses on modelling additive genetic random regression RN models, although the concepts are relevant to all types of RN models.

There are two types of $G \times E$ which can contribute to variation in the slope of a RN model, and both have very different properties. Rank-type $G \times E$ occurs when the genetic correlation between environments is less than unity. A correlation significantly less than unity means that between environments, genotypes re-rank and the relative effect of the underlying genes changes (Falconer, 1952). In contrast, scale-type $G \times E$ occurs when the genetic variance changes between environments. This “scales” the breeding values such that the differences between genotypes get larger or smaller, but it does not affect the genetic correlation (James, 2009). Separating the two sources of variation in the slope of a reaction norm would enable a clearer picture of re-ranking between genotypes, and how the genetic architecture of a trait

varies across the EC. If the effect of scale-type $G \times E$ was removed from the reaction norm slope, the resulting slope would capture how much the genotype re-ranks across the EC, or its “rank-stability”. This could assist breeders to select genotypes which combine high overall performance with a stable ranking across a given EC, or alternatively, genotypes with adaptation to a specific EC level.

A method suggested by Falconer (1990) proposes that for a linear RN model with the intercept placed at the average EC, the square of the correlation between the intercept and slope (r^2) represents the proportion of variance in the slope that is due to scale-type $G \times E$. As such, the proportion of variance in slope due to rank-type $G \times E$ can be estimated as $(1 - r^2)$. A genetic regression can also be used to derive breeding values for the reaction norm slope that are independent of the intercept variance (thus independent of scale-type $G \times E$), as recently applied in the reaction norm analysis of body weight in sheep (Waters et al., 2022).

A drawback of Falconer’s method is that it is limited to a linear function, which might not adequately describe the response of a genotype across the EC in some situations. This could be addressed by directly standardising the breeding value of genotypes across the EC to a constant genetic variance, thereby removing variation in stability due to scale-type $G \times E$. For a given genotype, the change in standardised breeding value could then be used as an estimate of stability due to rank-type $G \times E$.

Factor analytic (FA) models can also separate scale-type and rank-type $G \times E$ when analysing the stability of genotypes across environments (Smith & Cullis, 2018). While RN models employ a single user-defined environmental covariable, FA models generate an unsupervised set of c common factors that represent unobserved environmental covariables which explain changes in genetic effects across environments (Meyer, 2009). The slopes of regression on the common factors can be used to derive breeding values for stability while accounting for scale-

type $G \times E$ (Smith & Cullis, 2018). A downside of using FA models is that they reveal unobserved environmental factors which may not always be interpretable. Therefore, the model is less useful than RN models when investigating a specific environmental factor, such as in designed experiments. However, the ability of FA models to separate scale-type and rank-type $G \times E$ could be leveraged to investigate scale-correction methods for RN models.

The aim of this study was to apply the two methods for separating variation due to scale and rank-type $G \times E$ when estimating breeding values for stability from random regression RN models. The dataset involves a MET from an Australian barley breeding program with large amounts of scale-type and rank-type $G \times E$. The resulting breeding values for stability from the RN models were compared with those derived from FA models to test how well they accounted for scale-type $G \times E$. We use genomic data to model the additive genetic relationships among genotypes.

4.3 Methods

4.3.1 Phenotypic data

The data were provided by InterGrain Pty Ltd (www.intergrain.com) and consisted of plot yields from 52 partially replicated trials distributed across 15 locations throughout Western Australia, New South Wales, Victoria, and South Australia. The trials took place in 2019 (10 trials) and 2020 (42 trials).

There were 3460 unique genotypes with yield phenotypes, although 2734 of the genotypes were only represented in one environment and did not have single nucleotide polymorphism (SNP) data. The remaining 726 genotypes were almost fully replicated across the 15 environments and had SNP data.

4.3.2 SNP data

The SNP data was collected using the Illumina Infinium Wheat Barley 40K SNP array Version 1.0 and was imputed to a set of 410 499 SNPs, as described in Keeble-Gagnère et al. (2021). After imputation, SNPs with a minor allele frequency less than 0.05 (103 489 SNPs) and with more than 12.5% heterozygosity (202 SNPs) were removed, leaving a total of 306 808 SNPs for the 726 genotypes.

4.3.3 Analysis

Because there was a large number of genotypes without replication across environments, a two-stage analysis was used (Smith et al., 2001) rather than a usually preferable single stage. This allowed all the genotypes to be used for modelling spatial variation within the trials in the first stage, while avoiding the inclusion of 2734 uninformative genotypes to model G×E across environments in the second stage. All the models were fit using restricted maximum likelihood (REML) in ASReml-R (Butler et al., 2018).

In the first stage, trials were analysed independently to obtain best linear unbiased estimates (BLUEs) of yield for each genotype, along with the corresponding weights. For each trial, the initial model consisted of an independent random genotype effect and a first-order separable autoregressive term for the error variance structure using the row and column dimensions of the trial. Based on this model, plots with mean yields deviating more than four standard deviations from their expectation were removed. Additional random ‘row’ and ‘column’ terms were fitted based on a log-likelihood ratio test with a threshold of $p = 0.05$. Once the final model was determined, the BLUEs and their corresponding weights were obtained by fitting genotype as a fixed effect, following the methodology of Smith et al. (2001).

In the second stage, the performance of genotypes across the environments was analysed with RN and FA models utilizing the BLUEs of yield and corresponding weights obtained in the

first-stage analysis. Only the 726 genotypes with replication across environments and SNP data was considered in this stage. Of the 726 genotypes, 74 were removed from the analysis, as they were heterozygous for more than 12.5% of their SNPs. This left 652 genotypes, of which 606 were replicated across all 15 environments, with an additional 45 genotypes replicated in 14 environments. The remaining genotype was replicated in only six environments but was included in the analysis since it would contribute information through its genomic relationship to the population.

4.3.4 Reaction norm models

The RN models require an environmental covariable (EC) that captures the environmental quality. While it is increasingly common for ECs to consist of real environmental covariables, this was not available in our dataset. Instead, a joint-regression analysis was used (Eberhart & Russell, 1966; Finlay & Wilkinson, 1963; Perkins & Jinks, 1968), where the adjusted mean yield of each environment was used as the EC. RN models based on joint-regression analysis have been used widely to investigate the genetic basis of stability for yield and quality traits in cereal crops (Calderini & Slafer, 1999; Kraakman et al., 2004; Li et al., 2019; Sandhu et al., 2021; Tan et al., 2020). Because the distribution of genotypes across environments was slightly unbalanced in our dataset, the direct mean of yield in each environment would be a slightly biased estimator of the EC. To account for the unequal use of genotypes across environments, a simple model which fit genotypes as an independent random effect and environment as a fixed effect was used to obtain a BLUE of yield for each environment, which formed the EC. The EC values ranged from 1.45 to 4.84 tonnes per hectare (t/ha) and were standardised to values with a variance of one and a mean of zero for the RN analysis. The distribution of genotypes along the EC is given in Appendix B, Figure B4.1.

The first RN model was a standard linear reaction norm (RN-L), which modelled the genetic effect of each genotype as a linear function of the EC. The second RN model used cubic

polynomials (i.e., a non-linear function) to model the genetic effect across the EC and was denoted RN-NL.

The linear RN-L in simple matrix form was:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a}_0 + \mathbf{Z}_2\mathbf{a}_1 + \mathbf{e} \quad (1)$$

Where $\mathbf{y}$ was a vector of BLUEs for each genotype, $\mathbf{X}$ was an incidence matrix linking records to the fixed effects $\mathbf{b}$, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ were design matrices linking records to the genomic estimated breeding values (GEBVs) for the intercept ($\mathbf{a}_0$) and slope ($\mathbf{a}_1$), with $\mathbf{Z}_1$ containing 1's on the diagonal and $\mathbf{Z}_2$ containing the EC value corresponding to each BLUE in $\mathbf{y}$ on the diagonal, and $\mathbf{e}$ was the residual variance.

The environment was fitted as the only fixed effect to account for differences in the mean yield between environments, and residual variance was estimated independently for each environment. The genetic variance of $\mathbf{a}_0$ and $\mathbf{a}_1$ was modelled according to:

$$\begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \end{bmatrix} \sim N(0, \mathbf{G} \otimes \mathbf{K}), \text{ where } \mathbf{K} = \begin{bmatrix} \sigma_{a_0}^2 & \sigma_{a_1 a_0} \\ \sigma_{a_0 a_1} & \sigma_{a_1}^2 \end{bmatrix} \text{ and } \mathbf{G} \text{ was a relationship matrix based on SNP}$$

data, constructed using the first van Raden method (VanRaden, 2008) and $\otimes$ refers to a Kronecker product of matrices.

For RN-NL, a cubic polynomial was fitted to describe the genetic effects across the EC. The model was a simple extension of RN-L:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a}_0 + \mathbf{Z}_2\mathbf{a}_1 + \mathbf{Z}_3\mathbf{a}_2 + \mathbf{Z}_4\mathbf{a}_3 + \mathbf{e} \quad (2)$$

Where $\mathbf{Z}_3$ and $\mathbf{Z}_4$ were design matrices linking records to the GEBVs for quadratic ($\mathbf{a}_2$) and cubic ($\mathbf{a}_3$) components, with the diagonal of $\mathbf{Z}_3$ and $\mathbf{Z}_4$ containing the square and cube of the EC value corresponding to the BLUE in $\mathbf{y}$ respectively. The remaining components $\mathbf{y}$, $\mathbf{X}$, $\mathbf{b}$,

$\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{a}_0, \mathbf{a}_1$ and $\mathbf{e}$ were the same as in (1). The genetic variance of $\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2$ and $\mathbf{a}_3$ were also modelled in the same way as in RN-L, where:

$$\begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix} \sim N(0, \mathbf{G} \otimes \mathbf{K}), \text{ where } \mathbf{K} = \begin{bmatrix} \sigma_{a_0}^2 & \sigma_{a_1 a_0} & \sigma_{a_2 a_0} & \sigma_{a_3 a_0} \\ \sigma_{a_0 a_1} & \sigma_{a_1}^2 & \sigma_{a_2 a_1} & \sigma_{a_3 a_1} \\ \sigma_{a_0 a_2} & \sigma_{a_1 a_2} & \sigma_{a_2}^2 & \sigma_{a_3 a_2} \\ \sigma_{a_0 a_3} & \sigma_{a_1 a_3} & \sigma_{a_2 a_3} & \sigma_{a_3}^2 \end{bmatrix}$$

For both RN models, the genetic (co)variance matrix ($\mathbf{A}$) across the EC was obtaining using:

$$\mathbf{A} = \boldsymbol{\Phi} \mathbf{K} \boldsymbol{\Phi}' \quad (3)$$

In RN-L, $\boldsymbol{\Phi}$ was a 15×2 matrix containing a vector of 1's in the first column and a vector of the EC values for each environment. In RN-NL, $\boldsymbol{\Phi}$ was a had an additional two columns (15×4) consisting of the EC values squared and cubed respectively.

4.3.5 RN overall breeding value and stability estimates

In both RN-L and RN-NL, the GEBV for the intercept ($\mathbf{a}_0$) was treated as the overall GEBV for each genotype. The slope GEBV ($\mathbf{a}_1$) in RN-L captured how much the overall GEBV of each genotype changed across the EC and was therefore used as a GEBV for stability. This is the standard model for deriving stability from linear RN models (e.g., de Souza et al., 2020; Kusmec et al., 2017; Lacaze et al., 2009), so it was treated as the control measure of stability. None of the GEBVs in RN-NL could be directly interpreted as stability. Stability could only be estimated in RN-NL using the second scale-correction method (outlined later).

The slope GEBV ($\mathbf{a}_1$) could include variation due to both scale-type and rank-type $G \times E$. The following two methods aim to remove the variation due to scale-type $G \times E$, so that the resulting GEBVs describe variation due to rank-type $G \times E$ across the EC.

The first method was derived from Falconer (1990). Here, the correlation between the intercept and slope is attributed entirely to scale-type $G \times E$, the while the remaining variation that is

independent of the correlation is attributed to rank-type G×E. We can obtain GEBVs for the slope that are independent of the genetic correlation between the intercept and slope by using a genetic regression:

$$\mathbf{a}_1^* = \mathbf{a}_1 - \frac{\sigma_{a_1 a_0}}{\sigma_{a_0}^2} \mathbf{a}_0 \quad (4)$$

Where $\mathbf{a}_0$ and $\mathbf{a}_1$ are the GEBVs for the intercept and slope, respectively, $\sigma_{a_0}^2$ is the variance in intercept, and $\sigma_{a_1 a_0}$ is the covariance between the intercept and slope. The resulting GEBVs ($\mathbf{a}_1^*$) are the essentially the residuals of a genetic regression between the intercept and slope components. This method, referred to hereafter as the genetic regression, was only applied to RN-L, as it cannot be used for reaction norms with higher-order polynomials (i.e., RN-NL).F

The second method also attempts to remove variation in the reaction norm slope arising from scale-type G×E while allowing the use of RN models with higher-order polynomials. Briefly, the GEBVs of genotypes along the EC are standardised to a constant genetic variance to remove scale-type G×E. After standardisation, the change in GEBV for each genotype across the EC is used to estimate stability. This method, hereafter referred to as the standardisation method, was applied independently to both RN-L and RN-NL.

In more detail, the GEBV for genotype i at EC value t in a RN model is a function of the GEBVs for each regression coefficient ($a_0 \dots a_p$) and was calculated as:

$$\text{GEBV}_{it} = a_{0i} + a_{1i} \times t + \dots + a_{pi} \times t^p$$

Likewise, the genetic variance at the different values of t can be calculated using $\mathbf{A} = \boldsymbol{\Phi} \mathbf{K} \boldsymbol{\Phi}'$, as described in Eq. 3. The GEBV for each genotype at each value of t can then be standardised (SGEBV) based on the genetic variance.:

$$\text{SGEBV}_{it} = \frac{\text{GEBV}_{it}}{\sqrt{\text{Va}_t}}$$

Where Va_t is the genetic variance at EC value t . By performing this across the range of EC values in the data, the SGEbVs for each genotype can be plotted across the EC. The resulting standardised reaction norm describes the breeding value of each genotype across the EC whilst being adjusted for changes in the scale of the genetic variance. The SGEbVs only represent relative performance and do not carry units. Since the SGEbVs are not linear with respect to t , there is not a simple measure for stability. A solution is to use the change in SGEbV over a specified interval across the EC to estimate the stability (**S**):

$$S_i = \frac{SGEBV_{i,t_x} - SGEbV_{i,t_y}}{t_x - t_y} \quad (5)$$

Where t_x and t_y represent the maximum and minimum EC values in the specified interval. It is therefore possible to estimate the stability of each genotype for different segments of the EC. This could be used to explore whether different genotypes are more sensitive to different parts of the EC. The current study used an interval containing the entire range of the EC. Therefore, **S** was interpreted for each genotype as the change in GEBV relative to the population across the range of the EC and was calculated using both RN-L and RN-NL.

4.3.6 Factor analytic models

Unlike RN models, FA models ‘find’ environmental variables, called common factors, that explain patterns of $G \times E$ across environments. The regression on these common factors can be used to estimate the stability of genotypes across environments. If the EC used in the RN models is a major driver of $G \times E$, we can expect estimates of stability and overall performance to be similar between these methods. Since FA models have special properties which allow the user to measure the amount of scale-type $G \times E$ in the regression coefficients for the environmental variables, we used stability estimated from FA models as a reference to compare the effectiveness of the two RN correction methods to account for scale-type $G \times E$.

The first FA model (FA-2) used two common factors to model genetic effects that were common across the environments, along with specific genetic and residual variances for each of the 15 environments. FA-2 was described in matrix notation as:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_g(\mathbf{\Lambda}_a \otimes \mathbf{I}_n)\mathbf{f}_a + \mathbf{Z}_g\boldsymbol{\delta}_a + \mathbf{e} \quad (6)$$

Where $\mathbf{y}$, $\mathbf{X}$, and $\mathbf{b}$ were the same as in Eq. 1, $\mathbf{Z}_g$ was a design matrix linking records to the random effects, $\mathbf{\Lambda}_a = [\boldsymbol{\lambda}_{a1}, \boldsymbol{\lambda}_{a2}]$ was a $b \times 2$ matrix of estimated loadings for the b environments and first ($\boldsymbol{\lambda}_{a1}$) and second ($\boldsymbol{\lambda}_{a2}$) common factors, $\mathbf{I}_n$ was an $n \times n$ identity matrix, where n was the number of genotypes, $\mathbf{f}_a = (\mathbf{f}'_{a1}, \mathbf{f}'_{a2})'$ was a $n \times 2$ vector of genotype regression coefficients for the first ($\mathbf{f}_{a1}$) and second ($\mathbf{f}_{a2}$) common factor, $\boldsymbol{\delta}_a = (\boldsymbol{\delta}'_{a1}, \boldsymbol{\delta}'_{a2}, \dots, \boldsymbol{\delta}'_{ab})'$ was a $n \times b$ vector containing the specific genetic effects ($\boldsymbol{\delta}_{aj}$) for the b environments, and $\mathbf{e}$ was the residual variance. The residual variance was estimated independently for each environment.

The variance of the common factors and specific effects were modelled according to:

$$\begin{bmatrix} \mathbf{f}_a \\ \boldsymbol{\delta}_a \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{I} & \otimes \mathbf{G} & 0 \\ 0 & \boldsymbol{\Psi}_a & \otimes \mathbf{G} \end{bmatrix} \right)$$

where $\mathbf{I}_{c_a}$ was an identity matrix of order two, $\mathbf{G}$ was the genomic relationship matrix as calculated for the RN models, and $\boldsymbol{\Psi}_a$ was a $n \times n$ diagonal matrix with elements ψ_{aj} , which were the specific genetic variances for individual environments.

The matrix $\mathbf{A}$, which describes the additive genetic variance and covariance between the b environments, was calculated following:

$$\mathbf{A} = \mathbf{\Lambda}_a \mathbf{\Lambda}'_a + \boldsymbol{\Psi}_a$$

The second FA model (RR-2) was a reduced rank implementation of FA-2. RR-2 was identical to FA-2, except that the specific genetic variances for each of the 15 environments (ψ_{aj}) were

assumed to be zero. RR-2 was fitted because RN models also implicitly assume that the specific genetic effects are zero. Therefore, the results from RR-2 should account for the effect of estimating specific genetic effects when comparing the RN and FA models.

FA models with more than two common factors were not considered. This was because RN models essentially capture the performance of a genotype in two components: intercept (overall breeding value) and slope (change in breeding value over a unit change in EC). Similarly, the first common factor in a FA model usually represents the overall breeding value (comparable to the intercept), with the second common factor capturing stability (comparable to the slope). Going beyond two common factors would make it difficult to compare stability between FA and RN models and was not required to achieve the objective of the study, which was to test how well the RN model scale-corrections accounted for scale-type $G \times E$. For reference, the two common factors accounted for on average 72.8% of the additive genetic variance (see results section), while fitting a third common factor explained an additional 9.14% (results not shown).

4.3.7 Overall breeding value and stability in FA models

Breeding values for overall performance and stability for each genotype were derived from FA-2 and RR-2 based on the methods outlined in Smith & Cullis (2018). This involved first rotating the loadings and regression coefficients for each common factor using singular value decomposition to make them orthogonal. After rotation, the loadings were analysed to ascertain how much scale or rank-type $G \times E$ contributed to the regression coefficients for each common factor.

The sign of the loadings represents contrasts between the environments (Smith & Cullis, 2018). Since all genotypes have the same intercept for each common factor (0,0), there is no reranking of genotypes (or no contrast) across a common factor when the loadings all have the same sign. In this case, the regression coefficients will exclusively capture scale-type $G \times E$.

When there is a mix of positive and negative loadings, the regression coefficients capture at least some variation due to rank-type $G \times E$, because the genotypes pass through (0,0) and re-ranking occurs. The degree to which the regression coefficients represent rank-type $G \times E$ can be assessed by comparing the absolute mean of the positive and negative loadings (or average contrasts) respectively for the common factor. When the average contrasts are very different (Figure 4.1a), most of the variation in the regression coefficients will be due to scale-type $G \times E$, because the variance of GEBVs for the common factor will be very different between the average contrasts. When the absolute mean of the positive and negative loadings are equal, the variance of GEBVs for the common factor at the average contrasts will also be equal (Figure 4.1b), so the variation in regression coefficients will be mostly due to rank-type $G \times E$.

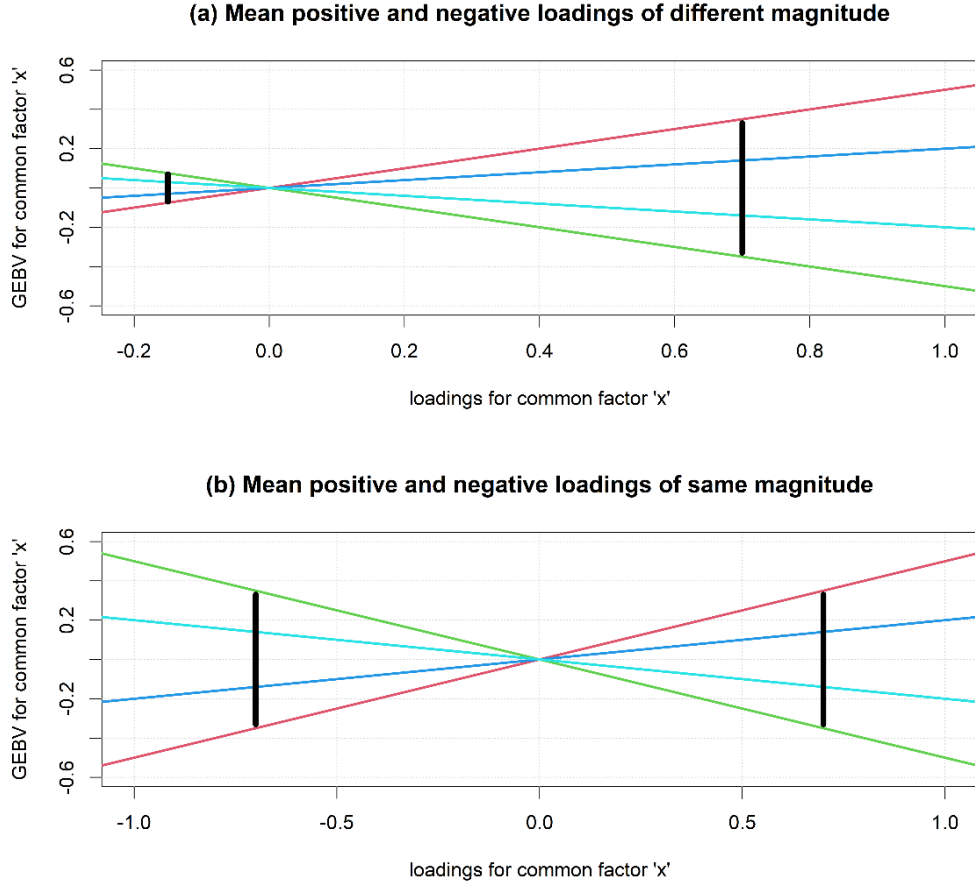


Figure 4. 1 Schematic example of the regression of four genotypes (coloured lines) across a common factor, 'x' in a FA model. The vertical black lines represent the variation in GEBVs for the common factor 'x' at the average of the positive and negative loadings (or average contrasts). In **a**, the average negative and positive loading are -0.15 and 0.7 units, respectively, and in **b**, -0.7 and 0.7 units respectively

According to Smith & Cullis (2018), the loadings for the first common factor are usually the same sign, while the remaining common factors tend to be a mix of positive and negative loadings. This experience was reflected in our analysis (see results). This meant the breeding value for overall performance across environments (**OP**) could be calculated as a function of the first common factor, since it represented a genetic effect common to all environments with no re-ranking. Therefore, **OP** was calculated as the mean additive genetic effect across the loadings for the first common factor:

$$OP_i = \frac{1}{b} \sum_{j=1}^b \tilde{\lambda}_{a_{1j}}^* \tilde{f}_{a_{1i}}^*$$

Where $\tilde{\lambda}_{a_{1j}}^*$ was the rotated loading for the first common factor for environment j , and $\tilde{f}_{a_{1i}}^*$ was the rotated regression coefficient for genotype i on the first common factor.

Because the second common factors in FA-2 and RR-2 contained a mix of positive and negative loadings that were similar in magnitude, they captured variation in stability due to rank-type $G \times E$. Therefore, the responsiveness of genotypes to the second common factor (**R2**) was used as the GEBV for stability and was calculated as:

$$R2_i = (\bar{\lambda}_{2+} - \bar{\lambda}_{2-})\tilde{f}_{a_{2i}}^*$$

Where $\bar{\lambda}_{2+}$ and $\bar{\lambda}_{2-}$ represent the mean of the rotated positive and negative loadings for the second common factor, respectively. A smaller value of **R2** indicated a genotype with greater stability in rank. The GEBVs for **R2** were used as the reference estimate of stability to which the various GEBVs for stability estimated from the RN models were compared.

The percentage of genetic variance in FA-2 that was explained by common factor d (either 1 or 2) in environment j in was calculated using:

$$v_{a_{dj}} = 100(\tilde{\lambda}_{a_{dj}}^*)^2 / \left(\sum_{s=1}^2 (\tilde{\lambda}_{a_{sj}}^*)^2 + \psi_{a_j} \right)$$

Where $\tilde{\lambda}_{a_{dj}}^*$ was the rotated loading for common factor d in environment j , and ψ_{a_j} was the specific genetic effect in environment j . The average value of v_{a_d} across the j environments was used to calculate the mean genetic variance explained by the common factors.

4.4 Results

4.4.1 Model summary

The FA models provided a better fit to the data than the RN models based on the Akaike information criterion (AIC), Bayesian information criterion (BIC) and log likelihood (Table 4.1). The linear RN model (RN-L) was a significantly poorer fit than the cubic RN model (RN-

NL) ($\Delta\text{AIC} = 139$), while the reduced-rank factor analytic model (RR-2) provided a poorer fit than the factor analytic model (FA-2) ($\Delta\text{AIC} = 1037$).

Overall, the first and second common factors in FA-2 explained 50.6 and 22.4% of the genetic variance across the environments, respectively (72.8% in total, Appendix B, Table B4.1). The percent of genetic variance explained by each common factor within the environments is also reported in Appendix B, Table B4.1. The genetic (co)variance of the random regression coefficients for RN-L and RN-NL are given in Appendix B (Table B4.2), along with correlations. The genetic correlation between the intercept and slope (r) in RN-L was high (0.76), indicating that genotypes with larger intercepts also tended to have larger slopes. Based on the r^2 , this also implied that 57.8% of the variation in slope was due to scale-type $G \times E$, while 42.2% was due to rank-type $G \times E$.

Despite the differences in model fit, the correlations between GEBVs for overall performance in the RN models (**a₀**) and FA models (**OP**) were high, ranging between 0.93 and 0.95 (Table 4.2). The lowest correlation between GEBVs for overall performance was between the RN models (0.90). Although not shown, the correlation between GEBVs for stability in FA-2 and RR-2 (**R2**) was also very high (0.96).

Table 4. 1 Summary of model fit. The number of parameters excludes those which fixed to zero during the model fitting process

^a Model	No. Parameters	^b AIC	^c BIC	Log Likelihood
RN-L	18	-8076.89	-7947.08	4056.445
RN-NL	25	-8215.64	-8035.35	4132.818
RR-2	44	-8565.06	-8247.76	4326.532
FA-2	58	-9601.79	-9183.52	4858.896

^aRN-L: linear reaction norm; RN-NL: non-linear reaction norm; RR-2: reduced-rank factor analytic model with two common factors and zero specific genetic effects; FA-2: factor-analytic model with two common factors and specific genetic effects

^b Akaike information criterion

^c Bayesian information criterion

Table 4. 2 Pearson correlations between GEBVs for overall performance in the RN models (**a₀**) and FA models (**OP**)

^a Model	RN-L	RN-NL	FA-2	RR-2
RN-L	1	0.90	0.93	0.95
RN-NL		1	0.94	0.95
FA-2			1	0.99

^aRN-L: linear reaction norm; RN-NL: non-linear reaction norm; FA-2: factor-analytic model with two common factors and specific genetic effects; RR-2: reduced-rank factor analytic model with two common factors and zero specific genetic effects

4.4.2 Magnitude of G×E

Genetic variance increased along the EC in RN-L and RN-NL by 430% and 486% respectively, indicating strong scale-type G×E (Figure 4.2). A smoothing spline fitted to the genetic variance of FA-2 and RR-2 also revealed similar increases in genetic variance across the EC. Overall, FA-2 estimated approximately 30-45% more genetic variance across the EC compared to RR-2, which was due to modelling specific genetic effects for each environment. The RN models estimated slightly less genetic variance across the EC compared to RR-2.

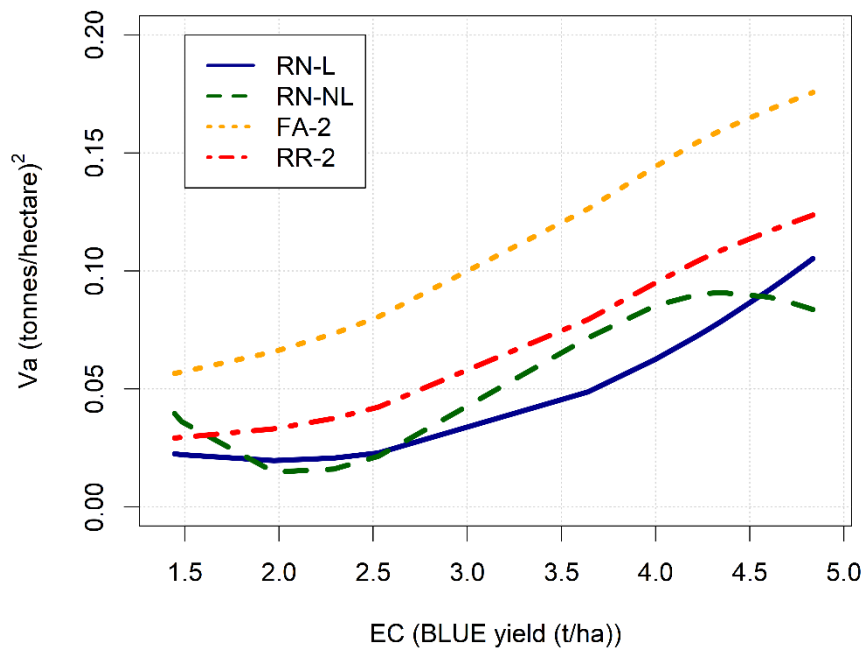


Figure 4. 2 Additive genetic variance (V_a) across the environmental covariable (EC) for the four models. In the RN-L and RN-NL, the line connects the estimates of V_a for each environment, while a smoothing spline was fitted to the estimates of V_a across the EC in RR-2 and FA-2

Genetic correlations between low and high yielding environments in the RN models were close to zero, indicating substantial levels of rank-type $G \times E$ (Table 4.3). The FA models tended to estimate lower genetic correlations between the environments.

Table 4. 3 Minimum (min), maximum (max) and mean pairwise correlations between the 15 environments for the four models. The mean correlation was calculated using Fisher's Z-transformation

^a Model	RN-L	RN-NL	FA-2	RR-2
Min	0.08	0.02	-0.43	-0.51
Max	1.00	1.00	0.97	1.00
Mean	0.96	0.89	0.50	0.86

^aRN-L: linear reaction norm; RN-NL: non-linear reaction norm; FA-2: factor-analytic model with two common factors and specific genetic effects; RR-2: reduced-rank factor analytic model with two common factors and zero specific genetic effects

4.4.3 Loadings of factor analytic models

The average of the negative loadings for the first common factor were 92.0 and 90.4% smaller than the average of the positive loadings in FA-2 and RR-2 respectively (Table 4.4). Additionally, only one loading in RR-2 and two loadings in FA-2 were negative for the first common factor (Appendix B, Table B4.3). Therefore, the first common factor captured variation that was almost entirely due to scale-type $G \times E$ in both models. The average of the negative loadings for the second common factor were only 34.1 and 1.9% smaller than the average of the positive loadings in FA-2 and RR-2, respectively. Because the magnitude of the average positive and negative loadings was similar (especially in RR-2), the second common factor captured variation primarily due to rank-type $G \times E$. Therefore, we can expect the regression on the second common factor, **R2**, to capture variation due to rank-type $G \times E$.

Table 4. 4 Absolute value of the mean REML estimates of the positive and negative loadings for the two common factors in FA-2 and RR-2.

^a Model	Common factor	Mean positive loading	Mean negative loading
FA-2	1	0.273	0.022
	2	0.126	0.083
RR-2	1	0.231	0.022
	2	0.105	0.103

^aRN-L: linear reaction norm; RN-NL: non-linear reaction norm; FA-2: factor-analytic model with two common factors and specific genetic effects; RR-2: reduced-rank factor analytic model with two common factors and zero specific genetic effects

4.4.4 Comparison of stability measures

The GEBVs for slope in RN-L ($\mathbf{a}_1$) were lowly correlated with $\mathbf{R2}$ estimated in both FA-2 (-0.24) and RR-2 (-0.26) (Figure 4.3a and b). Therefore, the standard method for determining stability in RN models gave very different results to the FA models. However, applying a genetic regression to the GEBVs (Eq.4) for slope ($\mathbf{a}_1^*$) more than tripled the correlation with $\mathbf{R2}$, to -0.80 in FA-2 and -0.85 in RR-2 (Figure 4.3c and d). Estimating stability after standardising the genetic variance in RN-L ($\mathbf{S}$) also increased the correlation with $\mathbf{R2}$ relative to $\mathbf{a}_1$, although the increase was more modest than $\mathbf{a}_1^*$ (Figure 4.3e and f). Using the same standardisation method with RN-NL ($\mathbf{S}$), which used cubic polynomials, did not increase the correlation with $\mathbf{R2}$ compared to RN-L (Figure 4.3g and h). Note that the sign of the correlations is arbitrary, as the loadings and regressions in FA-2 and RR-2 can be multiplied by +1 or -1 while retaining their meaning (Smith & Cullis, 2018). They will be reported as absolute values herein.

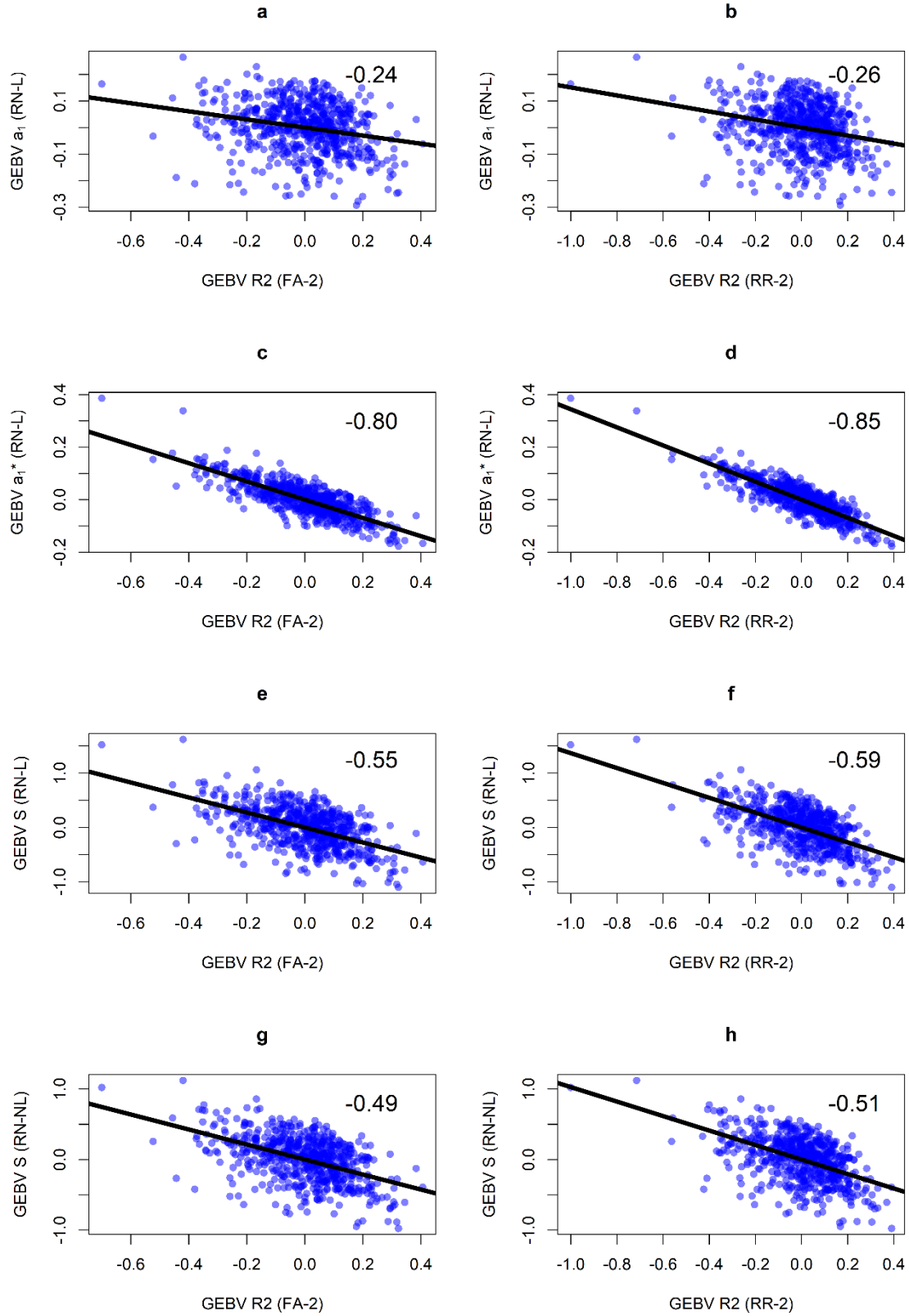


Figure 4. 3 Scatterplots between stability GEBVs from RN models (a_1 , a_1^* and S) compared with GEBVs for responsiveness to the second common factor (R_2) in FA-2 and RR-2. The correlations are printed in each frame. RN-L: linear reaction norm; RN-NL: non-linear reaction norm; FA-2: factor-analytic model with two common factors and specific genetic effects; RR-2: reduced-rank factor analytic model with two common factors and zero specific genetic effects

Relationship between overall performance and stability in different models

In RN-L, the genetic correlation between the intercept and slope (0.76, Appendix B, Table B4.2) and the correlation between GEBVs for the intercept ($\mathbf{a}_0$) and slope ($\mathbf{a}_1$) (0.73, Figure 4.4a) were high. This indicated that genotypes with a larger intercept tended to have larger change in GEBV across the EC and were therefore less stable. When the reaction norm slope was adjusted for scale-type G×E using the genetic regression ($\mathbf{a}_1^*$), the correlation with overall performance was close to zero (-0.08) (Figure 4.4b). This was also the case in FA-2 and RR-2 (Figure 4c and d). Although this was expected since $\mathbf{a}_1^*$ and $\mathbf{R}2$ were both estimated orthogonally to the overall performance, the correlation between $\mathbf{S}$ and $\mathbf{a}_0$ was also much smaller than 0.73 and made no assumption about the relationship between the intercept and slope (RN-L: 0.40, Figure 4.4e).

The correlation between $\mathbf{S}$ and $\mathbf{a}_0$ in the cubic RN-NL was comparably high (0.71, Figure 4.4f). However, the individual standardised reaction norms in this model were very noisy (see next section). This could cause the estimates of $\mathbf{S}$ (and therefore the correlation with $\mathbf{a}_0$) to be very sensitive to small changes in the interval which it was evaluated over (t_x and t_y), making interpretation of the correlation difficult.

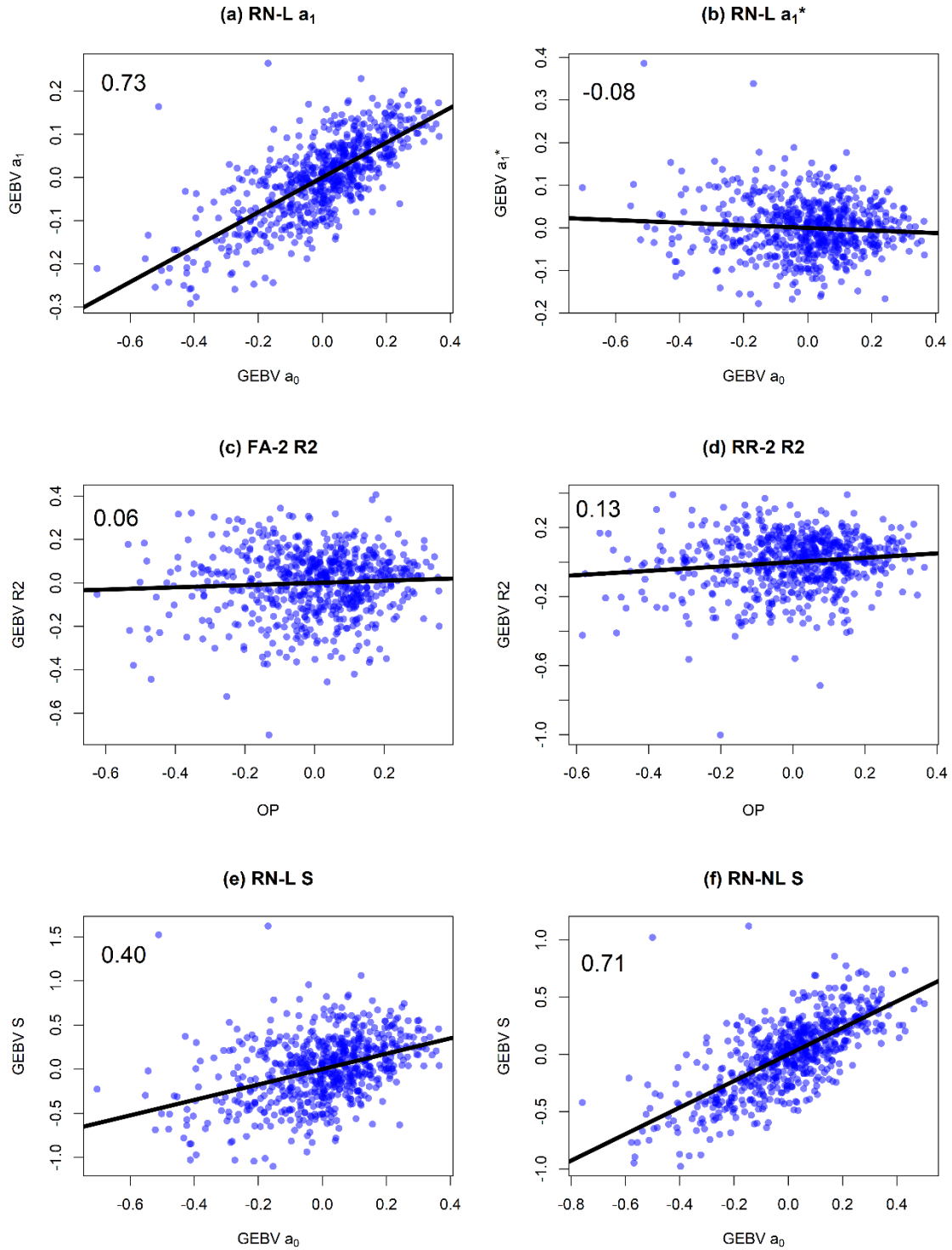


Figure 4.4 The relationships between overall performance (a_0 and OP) and stability (a_1 , a_1^* , S and $R2$) within the RN and FA models. The correlations are printed in each frame. RN-L: linear reaction norm; RN-NL: non-linear reaction norm; FA-2: factor-analytic model with two common factors and specific genetic effects; RR-2: reduced-rank factor analytic model with two common factors and zero specific genetic effects

Reaction norm scale corrections methods

To explore reaction norm scale corrections methods in more detail, the reaction norms for RN-L and RN-NL were plotted before and after applying the two correction methods (Figure 4.5). The two genotypes with very poor performance at low EC levels (coloured blue and black, Figure 4.5a, c) were ‘check’ varieties with lots of yield records and were not considered outliers in the analysis. Before the applying the correction methods, there was greater variation in GEBVs at higher mean yields and individual reaction norms crossed over, demonstrating both scale-type and rank-type $G \times E$ respectively (Figure 4.5a, c and f). When the slope of reaction norm was given by the genetic regression (a_1^*), the variation in GEBVs was more constant across the EC, although there was slightly more variation in the low-yielding environments (Fig 4.5b). This was likely because the mean EC (3.56 t/ha), which was the intercept, was slightly larger than the midpoint of the EC range (3.14 t/ha). The slope of regression changed direction for some genotypes.

When the GEBVs were standardised by the genetic variance, the reaction norms for RN-L were no longer linear, while the variation of GEBVs was also constant across the EC in both models (Figure 5d and g). This can be seen more clearly in Appendix B, Figure B4.2, which shows the variance of the GEBVs across the EC before and after standardisation. A linear approximation of the standardised reaction norms, which used the SGEBV when $EC = 0$ as the intercept and S as the slope, revealed that scale-type $G \times E$ was greatly reduced for RN-L (Figure 5e) but to a lesser extent for RN-NL (Figure 5h). The individual standardised reaction norms in RN-NL tended to be more erratic with very steep curves. This could indicate that the model was attempting to capture more complex $G \times E$ effects, rather than the general pattern across environments.

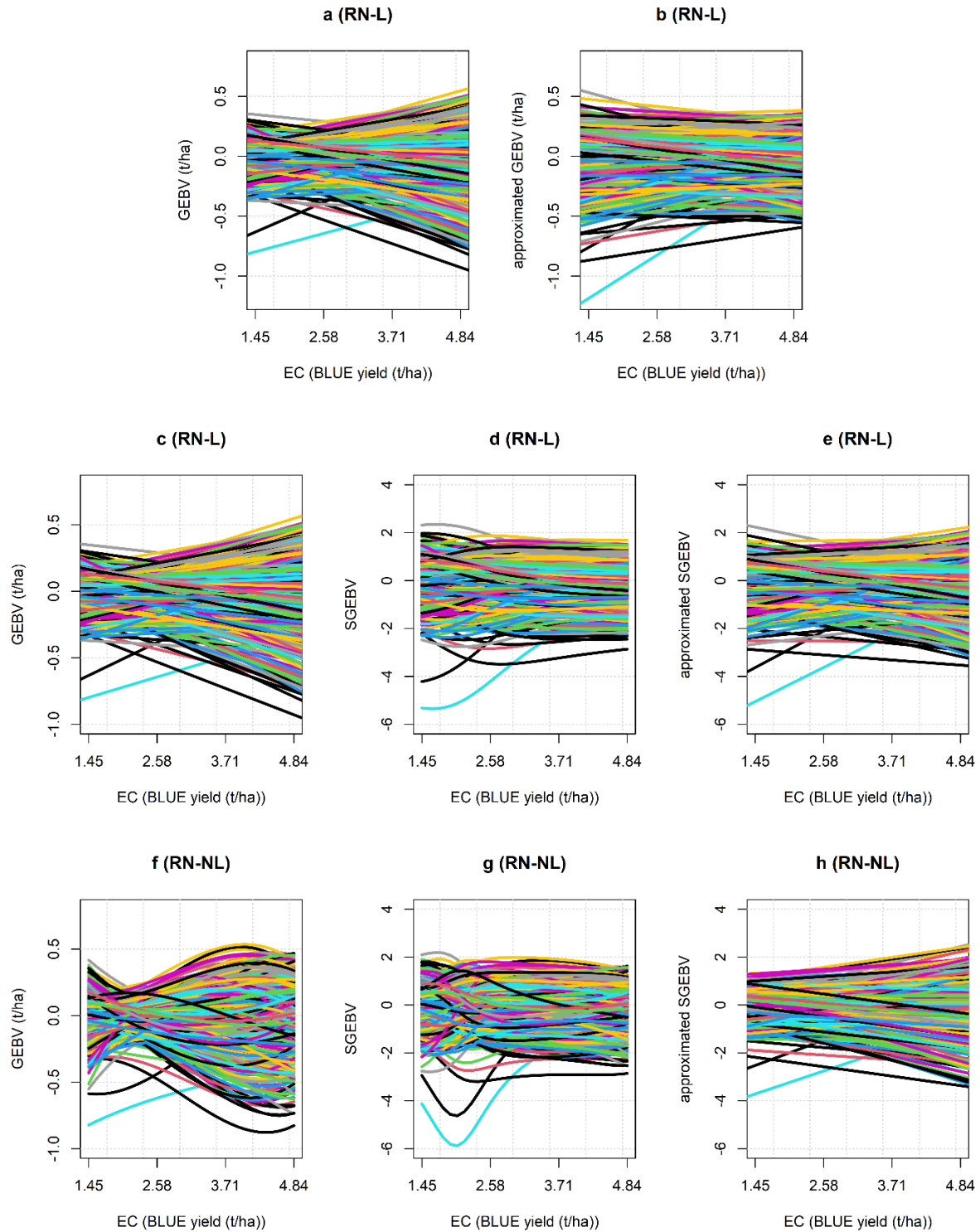


Figure 4.5 Reaction norms before and after applying the scale corrections. The top row shows the RN-L before (a) and after (b) applying the genetic regression, where the slope of each genotype is given by $\mathbf{a}_1$ and $\mathbf{a}_1^*$ respectively. The middle row shows RN-L before (c) and after (d) standardisation, and a linear approximation of the standardised reaction norm (e) where the slope of each genotype is given by their average change in SGEBV across the entire EC. The bottom row shows RN-NL before (f) and after (g) standardisation, and a linear approximation of the standardised reaction norm (h) where the slope of each genotype is given by their average change in SGEBV across the entire EC. RN-L: linear reaction norm; RN-NL: non-linear reaction norm

Discussion

The MET dataset contained large differences in the genetic variance (scale-type $G \times E$) and low genetic correlations (rank-type $G \times E$) between environments across the EC, which is expected in populations where $G \times E$ is significant. This level and type of $G \times E$ is typical of METs collected in Australian plant breeding programs (Smith et al., 2021). Hence, the data was a good resource for testing the scale-correction methods in this study.

The standard estimate of stability in a RN model is the breeding value for the slope. The slope GEBVs ($\mathbf{a}_1$) were very lowly correlated (0.24 to 0.26) with stability GEBVs estimated in the FA models ($\mathbf{R2}$) (Figure 4.3). This is interesting, as the RN and FA models would produce vastly different rankings of genotypes based on stability. Likewise, the SNP effects that underly the GEBVs would also be very different, which could lead to conflicting accounts of the genetic architecture of stability. However, correcting the slope GEBVs using a genetic regression ($\mathbf{a}_1^*$) tripled the correlation with $\mathbf{R2}$ (0.80 to 0.85). Because $\mathbf{R2}$ captured mainly rank-type $G \times E$ (Table 4.4), we can conclude that 1) $\mathbf{a}_1$ was lowly correlated with $\mathbf{R2}$ because it contained variation due to scale-type $G \times E$ and 2) $\mathbf{a}_1^*$ was highly correlated with $\mathbf{R2}$ because the variation due scale-type $G \times E$ in $\mathbf{a}_1$ was successfully removed by the genetic regression.

The GEBVs estimated using the standardisation method ($\mathbf{S}$) were also more highly correlated with $\mathbf{R2}$ compared to $\mathbf{a}_1$, although the increase was less than with $\mathbf{a}_1^*$. Therefore, it appears that the genetic regression is a more suitable approach for accounting for scale-type $G \times E$ in linear reaction norms. The correlation with $\mathbf{R2}$ was higher when $\mathbf{S}$ was estimated using the linear RN model (RN-NL) compared to the cubic RN model (RN-L). However, we would expect the linear reaction norm to be more closely related to $\mathbf{R2}$ than a cubic reaction norm, because $\mathbf{R2}$ is also a linear regression. Therefore, it was difficult to evaluate the merit of estimating $\mathbf{S}$ in either RN-L or RN-NL based on these results. Nevertheless, the standardisation method should

still be useful to account for scale-type $G \times E$ in situations which require curvilinear reaction norms.

The intercept and slope GEBVs were highly correlated in RN-L (Figure 4.4a, Table B4.2), indicating that genotypes with a larger intercept tended change more in GEBV across the EC and were therefore less stable. However, the correlation was close to zero when scale-type $G \times E$ was accounted for using a genetic regression and factor analytic models (Figure 4.4b, c and d), indicating there was no relationship between overall performance and stability due to rank-type $G \times E$. This is similar to other analyses of stability which account for scale-type $G \times E$ (Smith & Cullis, 2018; Tolhurst et al., 2019). This highlights the importance of understanding how stability is defined when interpreting its relationship with overall performance in different studies.

The additional benefit gained by applying the scale corrections is that the remaining slope variation should represent rank-type $G \times E$, which arises when the genetic correlation between environments is less than unity. This could be used to explore how the genetic architecture of a trait varies across an EC. This phenomenon was highlighted in a reaction norm study in Australian sheep (Waters et al., 2022), where a seemingly pleiotropic quantitative trait loci affected both the intercept and slope but disappeared when a genetic regression was used. Instead, a new region which contained a group of genes previously associated with environmental sensitivity in livestock was identified. It would be interesting to see if similar reports of pleiotropic genes affecting overall performance and stability in plants using reaction norms (e.g., Li et al., 2019) also change if additional analysis to account for scale type $G \times E$ is used.

There are exciting prospects for the improvement of yield stability in breeding programs using the genomic models in this study. Genomic breeding values derived from a training population

such as routine METs could be used in combination with a speed-breeding program to fast-track the creation of stable genotypes (Watson et al., 2018). In MET data, stability GEBVs derived from factor-analytic models would be recommended over the reaction norm models due to their superior ability for capturing the multifactorial causes of $G \times E$ in these settings (Kelly et al., 2007). This was realized in our analysis based on the AIC, BIC and the log-likelihood. Additionally, FA-2 estimated between 30 and 45% more genetic variance within environments compared to the remaining models, indicating it could disentangle $G \times E$ more efficiently. Despite this, the GEBVs from both models were highly correlated for overall performance (0.93-0.95) and stability (0.80 to 0.85) after scale-type $G \times E$ was accounted for. This is important, considering that the linear reaction norm required approximately two minutes to converge, compared to approximately 1 hour 10 minutes (RR-2) and 6 hours 4 minutes (FA-2) for the factor analytic models. As the number of genotypes and environments in routine analysis grow, the value of using genomic reaction norms for their speed also could increase. Examples of reaction norm models being used to efficiently fit dense genomic analyses with hundreds of thousands of genotypes can be found in livestock and human studies (Carvalho et al., 2019; Lee & van der Werf, 2016; Ni et al., 2019).

It is increasingly common for METs to collect environmental information to understand the causal factors underlying $G \times E$ (Cooper et al., 2021). A common way to incorporate the large number of resulting ECs is through (co)variance structures which describe the relationships between the environments based on the ECs (Jarquín et al., 2014; Raffo et al., 2022). Because these models implicitly regress phenotypes on ECs, they have also been interpreted as reaction norm models. While these models have demonstrated increased predictive ability over models which lack ECs (Cuevas et al., 2017; Lopez-Cruz et al., 2015), it is unclear whether they can reveal the stability of genotypes across the ECs due to the implicit nature of their fit. If methods

to extract the regressions of genotypes across the ECs become available, the impact scale-type $G \times E$ should be explored when estimating stability.

Another promising example for modelling this type of data integrates measured ECs into factor-analytic models (Tolhurst et al., 2022). In this study, the ECs explained mostly rank-type $G \times E$, while the unobserved common factors explained scale-type $G \times E$. Therefore, genotypes could be ranked based on their stability to the important ECs in the model. These models also outperformed equivalent reaction norms, which fit implicit random regressions on the individual environmental factors. Despite this, reaction norms should be the most efficient method for analysing stability when the EC is known and controlled (e.g., Houshmandfar et al., 2019; Paschke et al., 2003; Sadras et al., 2012), or the main source of $G \times E$ for the trait (Li et al., 2018; Millet et al., 2019).

Conclusion

This study investigated the ability of two methods to partition variation in the slope of a reaction norm model into the two types of $G \times E$ (rank-type and scale-type) which underly it. Stability estimated from factor analytic models, which explicitly disentangle rank-type and scale-type $G \times E$, was used for comparison. The two methods substantially increased the correlation with stability estimated from the factor-analytic models, indicating that they removed variation in the reaction norm slope that originated from scale-type $G \times E$. The genetic regression method appeared to be the most effective, yielding estimates of stability that were highly correlated (0.80 - 0.85) to those estimated from the factor analytic models. Although the standardisation method had a more modest increase, it could still be useful in situations where curvilinear RN models are required. Analyses which use reaction norms should consider implementing the scale-corrections outlined in this study to gain additional understanding of the nature of $G \times E$ and stability in their population.

Journal-Article Format for PhD Theses at the University of New England

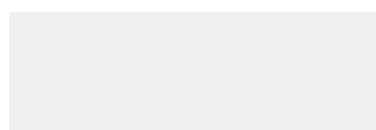
STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidates' contribution as indicated in the *Statement of Originality*.

	Author's Name	% of contribution
Candidate	Dominic Luke Waters	80
Other Authors	Julius van der Werf	6
	Hannah Robinson	3
	Lee Hickey	2
	Sam Clark	9

Name of Candidate: Dominic Luke Waters

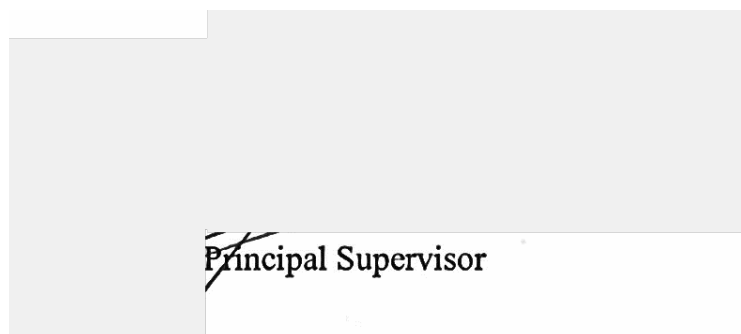
Name/title of Principal Supervisor: Professor Julius van der Werf



Candidate

30/06/2023

Date



Principal Supervisor

30/06/2023

Date

Journal-Article Format for PhD Theses at the University of New England

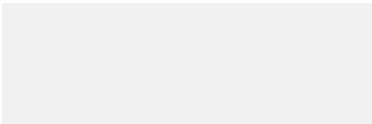
STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of Work	Page number's
Text	All
Figures	All
Tables	All

Name of Candidate: Dominic Luke Waters

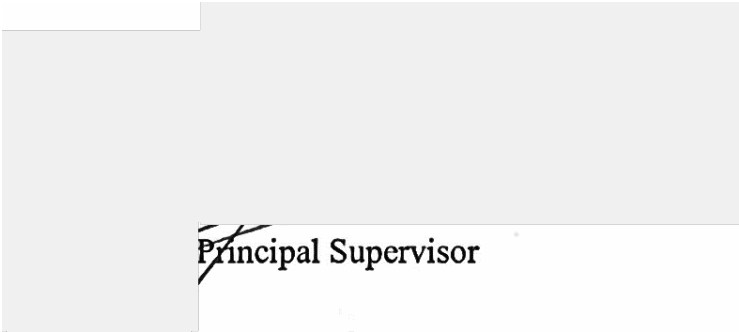
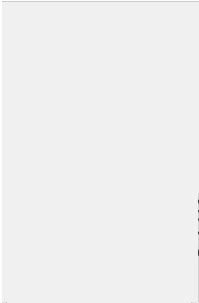
Name/title of Principal Supervisor: Professor Julius van der Werf



Candidate

30/06/2023

Date

Principal Supervisor

30/06/2023

Date

Chapter 5. Factor analytic models applied to the analysis of genotype by environment interactions and robustness in livestock

Dominic L. Waters^{1*}, Sam Clark¹, and Julius van der Werf¹

¹School of Environmental & Rural Science, University of New England, Armidale, NSW,
2351, Australia

*Corresponding author

E-mail addresses:

DLW: dwaters9@myune.edu.au

SAC: sam.clark@une.edu.au

JHVDW: jvanderw@une.edu.au

Prepared for Genetics Selection Evolution

5.1 Abstract

Background

Factor analytic models enable to estimation of genotype by environment (G×E) interactions between a large number of discrete environments. These models potentially offer a more flexible and less prescriptive way of capturing G×E interactions compared to the models normally applied to livestock data. The aim of this study was to investigate the use of factor analytic models for the analysis of G×E interactions in livestock data. Body weight records collected in a sheep research population with genetic linkage across flocks and years was used for the analysis. Reaction norm models, which are a popular choice for modelling were also fitted for comparison.

Results

A factor-analytic model with two common factors was sufficient to describe the G×E interactions between flock-years in the data. The model identified two clusters of environments in which performance was more correlated. The factor analytic ranking of genotypes for overall performance across environments was highly correlated with the reaction norm models (0.98) but was much lower for robustness to environmental variation (0.16-0.17). Relative to the reaction norms, the factor analytic models estimated higher average genetic variance (10.2 kg² vs 8.03 kg²), heritability (0.52 vs 0.42) and lower average genetic correlations (0.71 vs 0.96) between the flock-years.

Conclusions

Factor-analytic models can be used to explore G×E interactions and robustness in livestock data. The G×E interactions and estimates of robustness obtained from the factor analytic models were different to the more commonly applied reaction norm models. More research is

needed to understand the data structure required to accurately model G×E interactions with factor analytic models in livestock.

5.2 Background

Genotype by environment (G×E) interactions occur when the genetic correlation between a trait measured in different environments is less than one, or when the genetic variance of a trait changes between environments. These interactions can be accommodated in a genetic analysis by treating the performance in different environments as separate but correlated traits in a multiple trait model (Falconer, 1952; Mulder et al., 2004; Tsuruta et al., 2015; Williams et al., 2012). This results in a genetic (co)variance matrix, **G**, which describes the genetic variance for each environment, and the genetic covariance between environments. The number of genetic parameters required to fit **G** as an unstructured matrix increases exponentially with the number of environments ($p(p + 1)/2$, where p is the number of environments). Consequently, the estimation of **G** becomes computationally demanding when p is large, and the genetic parameters are estimated with less precision. This can be overcome by assuming the **G** matrix has an underlying structure that can be modelled using a smaller number of parameters.

In livestock, reaction norm models are often used to estimate **G** using a smaller number of parameters. These models assume that **G** can be described by the random regression of genetic effects across an environmental covariable (EC) that explains the G×E interactions. This covariable is often the performance of a contemporary group (e.g., Carvalho et al., 2019), or a combination of measured environmental factors (e.g., Nel et al., 2023). The number of genetic parameters to be estimated depends on the order of the polynomial used for regression; a linear regression requires only three genetic parameters (an intercept, slope, and a covariance between them). However, reaction norms could be limited by the strong assumptions placed on the importance of the EC. For instance, if the environment was described by the average

performance of cohorts, two contemporary groups could have the same performance for very different reasons (e.g., a lack of feed, or a high parasite burden). The genetic factors underlying performance in these environments would then be very different, but a reaction norm model would force them to have a genetic correlation of one and the same genetic variance. Likewise, when $G \times E$ interactions arise from many complex, interacting factors, a single composite EC might be overly simplistic and explain only a fraction of the total variance due to $G \times E$. Therefore, it could be useful to investigate other ways to parameterise $\mathbf{G}$.

Factor analytic (FA) models are used for the analysis of multi-environment trials in plant breeding, where genetic effects need to be accurately estimated in the presence of large and complex $G \times E$ interactions across many environments (Kelly et al., 2007). FA models approximate $\mathbf{G}$ by regressing genetic effects on several ‘latent’ common factors. The common factors represent unobserved environmental variables that are estimated from the data such that they to explain as much of the $G \times E$ variance between environments as possible (Cullis et al., 2014). For p environments and c common factors there are $p(c + 1) - c(c - 1)/2$ parameters to be estimated (Kelly et al., 2007), which represents a significant reduction compared to an unstructured covariance matrix when p is large. The common factors can be analysed to obtain useful information on the nature of $G \times E$ interactions acting in the population, such as clustering groups of environments in which there is minimal re-ranking of genotypes (Smith et al., 2021). The regressions can also be used to derive estimated breeding values (EBVs) for both the overall performance and robustness (referred to as ‘stability’ in plant breeding) of genotypes across environments (Smith & Cullis, 2018).

Although FA models have been suggested for the analysis of $G \times E$ interactions in livestock data (Meyer, 2009), there are limited examples of their use. Sae-Lim et al. (2014) used FA models to identify factors responsible for $G \times E$ interactions in rainbow trout. However, the small number of environments limited the model to a single common factor. This typically only

captures differences in genetic variance between environments, rather than differences in genetic correlations (Smith & Culls, 2018). Likewise, Li et al., (2019) modelled G×E interactions between different flocks of sheep using FA models but did not report details on the model and did not model the environmental effect of different years. Australian sheep are extensively managed across a diverse range of environments, and previous analyses have identified significant G×E interactions using reaction norm models (Pollott & Greeff, 2004; Waters et al., 2022). FA models present an interesting approach to analyse these interactions without making assumptions on the underlying environmental factors, as well as identify similar groups of environments and model variation in robustness.

The objective of this study was to analyse genotype by environment interactions in the body weight of sheep using factor analytic models and explore whether they identify G×E interactions differently to reaction norm models.

5.3 Materials and Methods

5.3.1 Dataset

Body weight records on 15,908 lambs raised in the Australian Sheep CRC Information Nucleus Flock (INF) were used for the analysis. The INF consisted of 8 flocks distributed across the diverse sheep-growing environments in Australia and was run for five years (2007-2011). Each year, approximately 100 sires were mated via artificial insemination at two of the main flocks (Armidale NSW and Katanning WA), with each flock containing 900 ewes. At the remaining six flocks, approximately 450 ewes were mated annually, using at least 50% of the Armidale/Katanning sires. Sires were generally used for one year across the flocks, while the dams were used across multiple years within the same flock. Since all lambs resulted from artificial insemination, the flock-year cohorts were all born within a three-week period. For

more information on the INF population, see van der Werf et al. (2010) and Waters et al. (2022).

Each flock-year was considered as a different environment. The genetic linkage between environments and the number of individuals per environment are important when fitting FA models (Lisle et al., 2021). The 31 flock-years had at least 350 lambs each, and the number of parents (sires, dams, grandsires and grandams; 7678 individuals in total) in common between flock-years is shown using a heatmap (Figure 5.1). On average, each flock-year shared 6.3% of the parents with any other given flock-year. However, genetic linkage tended to be stronger between flocks within years (due to common sires) and within flocks across years (due to common dams). Of the 465 pairwise flock-year combinations, 5 did not share any common parents. Although there are no published standards for genetic connectedness required to fit FA models accurately, the example detailed in Cullis et al. (2014) would indicate the level of genetic linkage in the current dataset is adequate. The genetic linkage was also stronger than indicated in Figure 5.1, as an additional 12,717 distant ancestors were available in the pedigree to calculate genetic relationships.

Pairwise Parental Connectivity

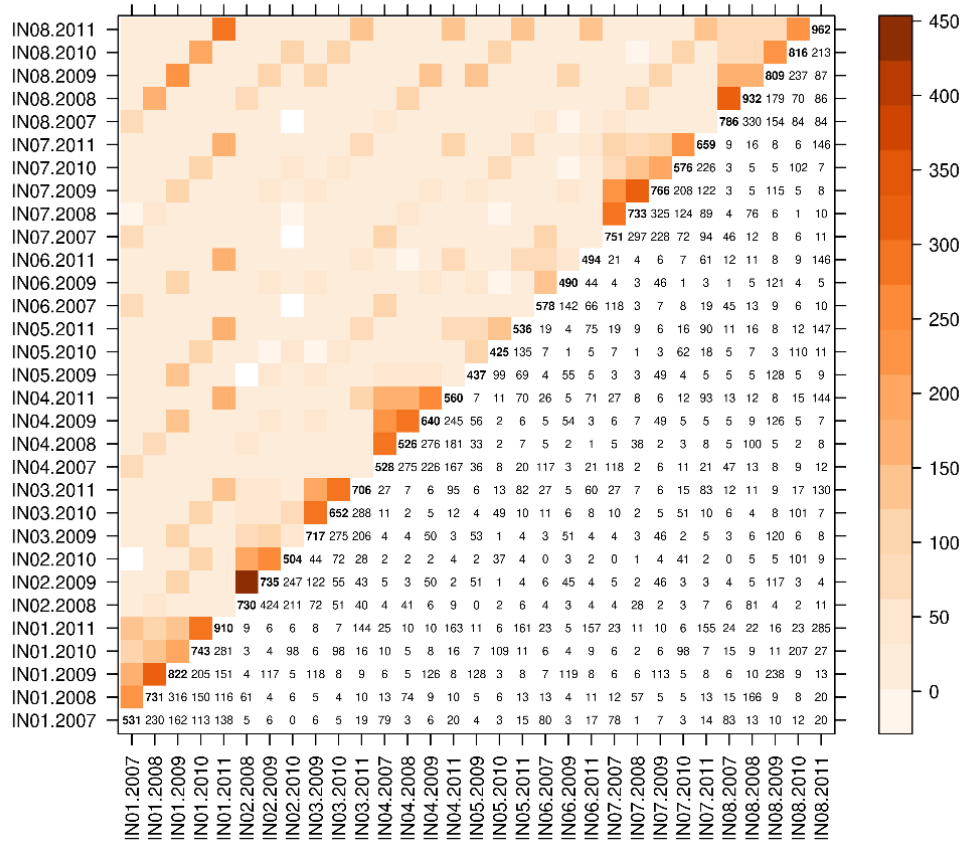


Figure 5. 1 The number of parents (sires, dams, grandsires and grandams) in common between the flock-years. The intensity of the heatmap colour indicates the number of parents, while the raw number is given below the diagonal. The number of parents for each flock-year is given on the diagonal in bold. The environments are ordered by year within flock (INxx.yyyy, where xx is the flock code and yyyy is the year)

Post-weaning weight (PWT), which was a bodyweight recorded between the ages of 129 and 239 days, was used for the analysis. The PWT phenotypes were pre-adjusted before the analysis, as there were difficulties in fitting the FA models with all the effects simultaneously.

The pre-adjusted phenotypes were obtained using an animal model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{c} + \mathbf{Q}\mathbf{g} + \mathbf{e} \quad (1)$$

Where $\mathbf{y}$ is the vector of PWT records, $\mathbf{X}$ is an incidence matrix for the fixed effects $\mathbf{b}$, $\mathbf{Z}_1$ was an incidence matrix relating the records to genetic effects ($\mathbf{a}$), $\mathbf{Z}_2$ is an incidence matrix relating records to the dam effects ($\mathbf{c}$), which were estimated ignoring relationships between dams, $\mathbf{Q}$ is a matrix of the proportion of each animal's genome originating from 39 breed-based genetic

groups, $\mathbf{g}$ is the vector of random genetic group effects, and $\mathbf{e}$ is the vector of residual genetic effects. Fixed effects included age at measurement (linear and quadratic), birth-type and rear-type interaction, sex, and contemporary group. The pre-corrected phenotypes ($\mathbf{y}^*$) were then calculated as:

$$\mathbf{y}^* = \mathbf{y} - \mathbf{Xb} - \mathbf{Z}_2\mathbf{c} - \mathbf{Qg} \quad (2)$$

A numerical summary of the data is given in Table 5.1.

Table 5. 1 Numerical summary of the dataset

Number of lambs	15,908
Number of sires of progeny with records	479
Number of dams of progeny with records	7609
Number of individuals in the pedigree	36,303
Number of contemporary groups	134
Contemporary groups per sire	8.4 ± 3.3
Contemporary groups per dam	1.7 ± 1.1
Number of flock-years	31
Flocks-years per sire of progeny with records	3.9 ± 1.3
Flock-years per dam of progeny with records	1.5 ± 0.8
Post-weaning age (days, mean $\pm$ standard deviation)	239.4 ± 51.4
Post-weaning weight (kg, mean $\pm$ standard deviation)	42.6 ± 8.7
Adjusted post-weaning weight (kg, mean $\pm$ standard deviation)	44.7 ± 4.6

5.3.2 Factor analytic models

To understand how FA models work, it is intuitive to first see how EBVs for each genotype in each environment are derived. The FA models consist of 1) a regression coefficient (f_a) for each genotype across each common factor, 2) the importance (or ‘loading’, λ_a) of each common factor in each environment, and 3) a residual genetic effect (δ_a) for each environment, which captures genetic variation within each environment that is not explained by the regression on the common factors. The EBV for genotype i in environment j is then given by:

$$EBV_{i,j} = \lambda_{a1j}f_{a1i} + \lambda_{a2j}f_{a2i} + \dots + \lambda_{acj}f_{aci} + \delta_{a1j} \quad (3)$$

Where λ_{a1j} is the loading of the d^{th} common factor for environment j , f_{a1i} is the regression coefficient of genotype i to the d^{th} common factor, and δ_{a1j} , is the residual genetic effects specific to environment j . Hence, environment j is characterised in the model by the loadings of the c common factors in that environment, plus the residual genetic effect. The performance of any genotype in environment j is then a function of its response to the c common factors, plus the residual genetic effect specific for environment j .

Factor analytic models with 1, 2 and 3 common factors (FA1, FA2 and FA3) were fit to the data. The models in matrix notation were:

$$\mathbf{y}^* = \mathbf{X}\mathbf{b} + \mathbf{Z}_g(\mathbf{\Lambda}_a \otimes \mathbf{I}_n)\mathbf{f}_a + \mathbf{Z}_g\mathbf{\delta}_a + \mathbf{e} \quad (4)$$

Where $\mathbf{y}^*$ was a vector of pre-adjusted PWT phenotypes, $\mathbf{X}$ was an incidence matrix linking records to the fixed effects $\mathbf{b}$ (i.e., the mean), $\mathbf{Z}_g$ was a design matrix linking records to the random effects, $\mathbf{\Lambda}_a = [\lambda_{a1}, \lambda_{a2}, \dots, \lambda_{ac}]$ was a $p \times c$ matrix of estimated loadings for the p environments and c common factors (λ_{ac}), $\mathbf{I}_n$ was an $n \times n$ identity matrix, where n was the number of genotypes in the pedigree ($n = 36,303$), $\mathbf{f}_a = (\mathbf{f}'_{a1}, \mathbf{f}'_{a2}, \dots, \mathbf{f}'_{ac})'$ was a $n \times c$ vector of

genotype regression coefficients for the c common factors, $\delta_a = (\delta'_{a1}, \delta'_{a2}, \dots, \delta'_{ap})'$ was a vector of length $n \times p$ containing the specific genetic effects for each environment, and $\mathbf{e}$ was the residual variance. The residual variance was estimated independently for each flock-year. The variance of the common factors and specific genetic effects were modelled according to:

$\begin{bmatrix} f_a \\ \delta_a \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{I}_{c_a} \otimes \mathbf{A} & 0 \\ 0 & \boldsymbol{\Psi}_a \otimes \mathbf{A} \end{bmatrix} \right)$, where $\mathbf{I}_{c_a}$ was an identity matrix of order c , $\mathbf{A}$ was the pedigree-based relationship matrix, and $\boldsymbol{\Psi}_a$ was a $p \times p$ diagonal matrix with elements ψ_{aj} , which were the residual genetic variances for individual environments. The matrix $\mathbf{G}$, which describes the genetic variance and covariance between the p environments and was used to derive the genetic correlation between flock-years, was calculated using:

$$\mathbf{G} = \mathbf{\Lambda}_a \mathbf{\Lambda}'_a + \boldsymbol{\Psi}_a \quad (5)$$

The heritability of PWT was calculated within environment j as the ratio of genetic variance in environment j to the sum of the genetic and residual variance in environment j . The models were fit using ASReml-R (Butler et al., 2018), which estimates the variance components using residual maximum likelihood.

5.3.3 Post analysis of factor-analytic models

To analyse the loadings and regression coefficients obtained from the FA models, they were rotated to a principal component solution using singular value decomposition, as described in Smith & Cullis (2018). This ensured that the common factors were orthogonal to each other and explained the maximum amount of genetic covariance between environments in the order of their fit (i.e., the first common factor explained the most genetic covariance between environments, the second common factor explained the next most, and so on).

Analysis of environments

The genetic variance explained by common factor d in environment j was calculated following Cullis et al. (2014):

$$v_{adj} = 100(\tilde{\lambda}_{adj}^*)^2 / \left(\sum_{s=1}^c (\tilde{\lambda}_{asj}^*)^2 + \psi_{aj} \right) \quad (6)$$

Where $\tilde{\lambda}_{adj}^*$ was the rotated loading for common factor d in environment j , and ψ_{aj} was the residual genetic effect in environment j .

Environments were clustered based on their similarity following the methodology of Smith et al. (2021). To describe briefly, the regression of genotypes along each common factor is estimated with the condition that all genotypes pass through the origin (0,0); that is, they all have the same intercept. As such, genotypes do not re-rank along a common factor if all of the loadings are of the same sign. By extension, we can create groups of environments that have the same signs for all the rotated common factors. The resulting groups will represent similar environments in which performance should be more highly correlated. The groups could then be interpreted using available agronomic, climatic, or other environmental knowledge to further understand the nature of G×E interaction in the population.

Calculation of overall performance and robustness

Estimated breeding values (EBVs) for the overall performance and robustness of each genotype were derived from the factor analytic models based on the methods outlined in Smith & Cullis (2018). The loadings for the first common factor were all positive values (see results) and represented a genetic effect common to all of the environments with no re-ranking. Therefore, the overall performance (**OP**) was calculated as the mean genetic effect across the loadings for the first common factor, which essentially is an EBV for PWT (kgs) at the average loading for the first common factor:

$$OP_i = \frac{1}{p} \sum_{j=1}^b \tilde{\lambda}_{a_{1j}}^* \tilde{f}_{a_{1i}}^* \quad (7)$$

Where $\tilde{\lambda}_{a_{1j}}^*$ was the rotated loading for the first common factor for environment j , and $\tilde{f}_{a_{1i}}^*$ was the rotated regression coefficient for genotype i on the first common factor. The responsiveness of genotypes to the second common factor (**R2**) was used as the EBV for robustness across environments and was calculated as:

$$R2_i = (\bar{\lambda}_{2+} - \bar{\lambda}_{2-}) \tilde{f}_{a_{2i}}^* \quad (8)$$

Where $\bar{\lambda}_{2+}$ and $\bar{\lambda}_{2-}$ represent the mean of the rotated positive and negative loadings for the second common factor, respectively. **R2** captured the change in PWT (kg) per unit change of the covariate implicit in the second common factor. A smaller value of **R2** indicated a genotype with greater robustness in ranking across environments. When more than two common factors are fitted, robustness can also be calculated as the residual mean square deviation (see Smith & Cullis, 2018). However, this was not calculated as the model with three common factors did not significantly improve the model fit.

5.3.4 Reaction norm models

The environmental covariable (EC) used to fit the reaction norm models was estimated as the adjusted post-weaning growth rate of each contemporary group, as described in Waters et al. (2022). The EC was considered at both the level of contemporary group and flock-year. In the case of flock-year, the EC was calculated as the average CG effect for each animal in the flock year. The reaction norm models in matrix form were:

$$\mathbf{y}^* = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a}_0 + \mathbf{Z}_2\mathbf{a}_1 + \mathbf{e} \quad (9)$$

Where $\mathbf{y}^*$, $\mathbf{X}$ and $\mathbf{b}$ were the same as in (4), $\mathbf{Z}_1$ and $\mathbf{Z}_2$ were design matrices linking records to the EBVs for the intercept ($\mathbf{a}_0$) and slope ($\mathbf{a}_1$), with $\mathbf{Z}_1$ containing 1's on the diagonal and $\mathbf{Z}_2$ containing the EC value corresponding to each record in $\mathbf{y}^*$ on the diagonal, and $\mathbf{e}$ was the

residual variance. The variance of the intercept and slope was modelled as: $\begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \end{bmatrix} \sim N(0, \mathbf{A} \otimes \mathbf{K})$

where $\mathbf{K} = \begin{bmatrix} \sigma_{a_0}^2 & \sigma_{a_1 a_0} \\ \sigma_{a_0 a_1} & \sigma_{a_1}^2 \end{bmatrix}$ and $\mathbf{A}$ is the pedigree relationship matrix.

Three different RN models were fitted using different ECs and residual variance structures. The first represented a typical linear RN model, where the EC consisted of the CG effects, and the residual was estimated across 5 discrete intervals along the EC (RN_CG). The second model RN model assumed that the environment was at the flock-year level, which was the same assumption used by the FA models. As such, the EC consisted of the flock-year effects (i.e., the average of the CG effects within each flock-year), and the residual variance was estimated independently for each flock-year (RN_FY). The third RN model used the CG effects as the EC but estimated the residual variance independently for each flock-year (RN_CG_FY). The genetic (co)variance matrix ($\mathbf{A}$) across the EC for all RN models was obtained using $\mathbf{G} = \mathbf{AKA}'$, where $\mathbf{A}$ contained two columns; the first was a vector of 1's, and the second was a vector of EC values. The heritability of PWT at a given EC level was obtained by dividing the genetic variance at the EC level by the sum of the genetic and residual variance at the EC level.

The EBVs for the intercept ($\mathbf{a}_0$) and slope ($\mathbf{a}_1$) were interpreted as EBVs for the overall performance and robustness of each genotype, respectively. Additionally, scale corrected EBVs for the slope were also estimated using a genetic regression, to capture variation in the slope that was independent of the genetic correlation between the intercept and slope:

$$\mathbf{a}_1^* = \mathbf{a}_1 - \frac{\sigma_{a_0 a_1}}{\sigma_{a_0}^2} \mathbf{a}_0 \quad (10)$$

The reaction norm models were also fitted using ASReml-R (Butler et al., 2018).

5.4 Results

5.4.1 Model fit

To select the optimal number of common factors in a factor analytic model, it has been recommended to examine the average percentage of genetic variation that is explained by the common factors across environments (Figure 5.2), in combination with formal model selection criteria (Smith et al., 2015; Tolhurst et al., 2019). A minimum of 80% genetic variance explained on average across the environments is often used to select the model (Cullis et al., 2014; Smith & Cullis, 2018; Tolhurst et al., 2019), in which case the FA2 model would be selected. This agreed with the formal model testing criteria (Table 5.2), which indicated the FA2 model provided the best fit for the data based on a log-likelihood ratio test and the AIC (Table 5.2). Therefore, the FA2 model was selected for the analysis of G×E interactions. Note that FA2 used 72 genetic parameters (excluding the residual variance) to estimate the genetic covariance between the 31 flock years, while an unstructured multi-trait model would have required 465 parameters.

Of the reaction norm models, RN_CG_FY provided the best fit across all the model testing criteria. This indicated that the model using contemporary group effects as an EC was superior at describing the G×E interactions relative to the model using flock-year effects. It also indicated the residual variance was best modelled individually for each flock-year, rather than the dividing it by groups based on the EC. The BIC, which penalises the number of parameters in a model more heavily than the AIC and log-likelihood ratio test, preferred the reaction norm models over the FA models.

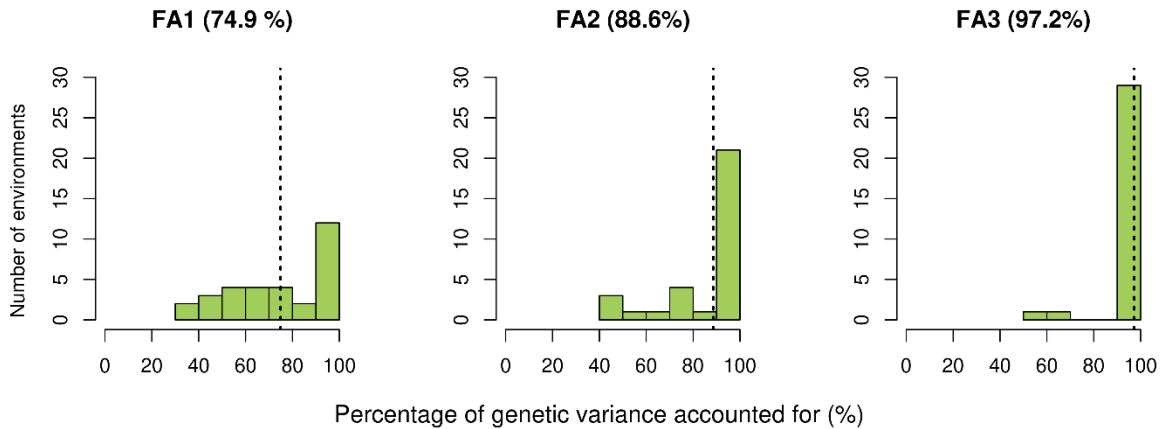


Figure 5. 2 Distribution of the percentage of genetic variance accounted for by the common factors in the FA models across the different environments (flock-years). The average percentage of genetic variation that is explained by the common factors across the environments is given in by the vertical dashed line and in brackets. FAX refers to a model with x common factors.

Table 5. 2 Summary of model fit for the different models. The values are reported as a deviation from the reaction norm model RN_CG. The number of parameters excludes those which were fixed to zero during the model fitting process. The model with the best fit for each criterion is highlighted in bold.

^a Model	Number of Parameters	Log Likelihood ^b	^c AIC	^d BIC
RN_CG	8	-31389.1	62794.3	62855.7
RN_FY	34	265.4	-478.9	-279.3
RN_CG_FY	34	269.8	-487.7	-288.2
FA1	86	343.8	-531.6	67.0
FA2	103	361.5	-533.1	196.0
FA3	124	376.5	-522.9	359.7

^aRN_CG: reaction norm with contemporary group EC, five discrete intervals for residual variance; RN_FY: reaction norm with flock-year EC, residual variance estimated for each flock-year; RN_CG_FY: reaction norm with contemporary group EC, residual variance estimated for each flock-year; FA1-FA3; factor analytic model with 1, 2 and 3 common factors, respectively.

^bA log likelihood ratio test with a significance threshold of $p = 0.05$ was used to test model fit

^c Akaike information criterion

^d Bayesian information criterion

5.4.2 G×E interaction analysis with FA2

The distribution of the rotated loadings across environments for the first and second common factor in FA2 are given in Figure 5.3. The first common factor explained on average 74.2% of the genetic variance across environments and were all positive. The variation between genotypes in the regression arose solely due to differences in the scale of the genetic variance, because the loadings were all the same sign (recall that each genotype has an intercept (0,0) for the regression across a common factor). As such, the regression on the first common factor represented of a large genetic effect that was consistent across all the 31 environments. Selection based on the first common factor should yield a positive response in post-weaning weight across all of the environments, although we would expect the response to be greater in environments with larger loadings for the first common factor.

In contrast, the loadings for the second common factor were distributed around zero. The regression represented G×E interaction due to re-ranking because there was a mix of positive and negative loadings. This means that selection based on the second common factor would cause a positive response in some environments and a negative response in others. However, the regression on the second common factor can also be used to identify genotypes with robust genetic merit across environments, as they will have a regression value close to zero.

The EBVs for overall performance and robustness, which were derived from the regression on the first and second common factors respectively, are plotted for a subset of 351 sires with at least 25 progeny in Figure 5.4. There was significant genetic variation in both the overall performance and robustness across environments that could be utilised through selection. The correlation between the EBVs was very low (-0.04), and sires that combined both high overall performance and stability (i.e., a large positive **OP** and small **R2**) can be identified in the plot.

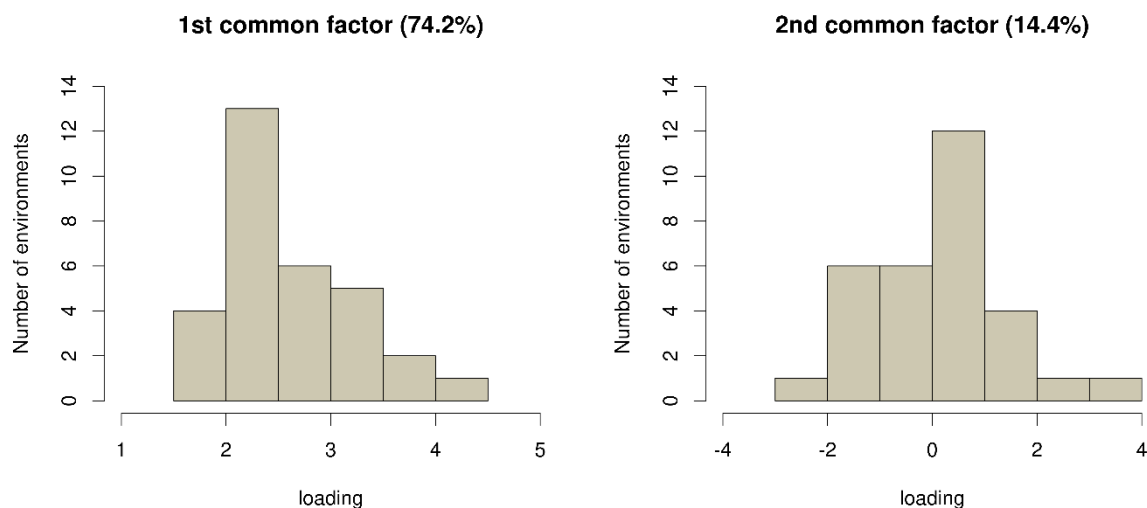


Figure 5.3 Distribution of the rotated loadings for the first and second common factor across environments (flock-years) in FA2. The percentage of genetic variance explained by each common factor is given in brackets.

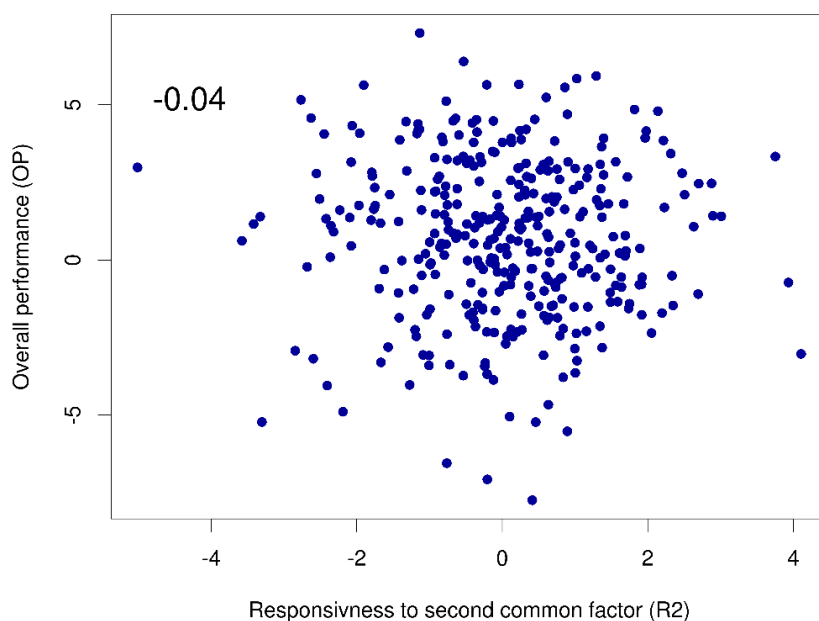


Figure 5.4 EBVs for overall performance (OP) and responsiveness to the second common factor (R2) for 351 sires with at least 25 progeny.

The genetic correlations between breeding values in the different flock-years ranged from -0.15 to 0.99 and averaged 0.72, indicating the presence of significant G×E interactions (Figure 5.5). Two groups of environments, delineated by the red lines, were clustered based on the sign of the loadings for the second common factor. The correlations within the first and second

group averaged 0.88 and 0.77, respectively, while the correlation between the two groups averaged 0.63. Of the 8 flocks, 3 occurred in only one group, which indicates that the environmental covariate implicit in the second common factor was consistent within these flocks. However, the remaining flocks were approximately equally distributed between the groups, indicating that the environments can differ significantly between years within flocks. There was also no obvious clustering of different years within the groups. The heritability of PWT varied significantly across environments, ranging from 0.26 to 0.79, with an overall mean of 0.52 (Figure 5.6).

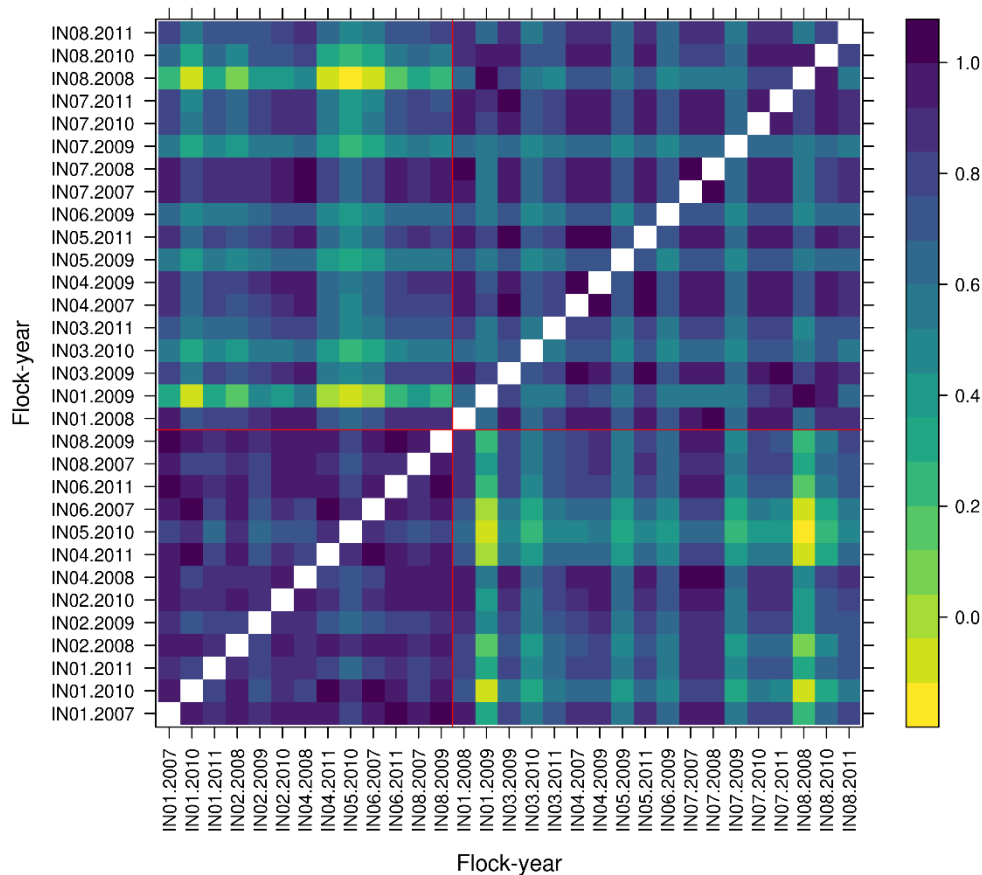


Figure 5. 5 A heatmap showing the genetic correlations between the 31 different environments (flock-years) in the data. The environments are divided into two classes based on their similarity using the interaction classes method (bottom left = class 1, top right = class 2). The boundary between the classes given by the red lines.

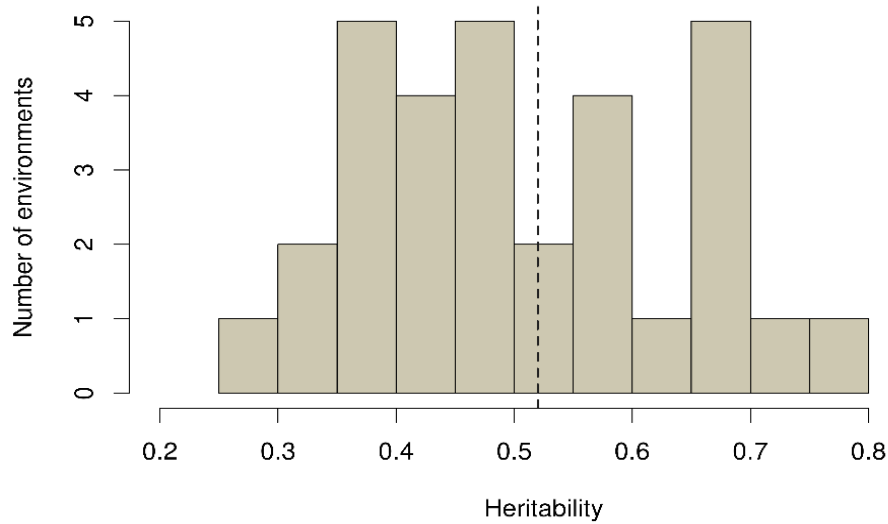


Figure 5. 6 Distribution of heritability across the environments in FA2. The mean (0.52) is given by the dashed vertical line

5.4.3 Comparison with reaction norm models

The reaction norm models all captured significant variation in the intercept and slope (Table 3). The RN_CG_FY model, which had the best fit of the reaction norms (Table 5.3), was used for comparison with the FA2 model. The average EC value within each flock-year was used to estimate the heritability, genetic variance and genetic correlations in RN_CG_FY.

Table 5. 3 Variance of the genetic intercept (σ_{a1}^2), slope (σ_{a2}^2), the ratio ($\sigma_{a2}^2 / \sigma_{a1}^2$) and the correlation between them (r_{a1a2}). The standard errors are given in brackets.

Model	σ_{a0}^2	σ_{a1}^2	r_{a1a2}
RN_CG	8.45 (0.51)	0.48 (0.17)	0.06 (0.10)
RN_FY	7.42 (0.47)	1.65 (0.43)	0.04 (0.09)
RN_CG_FY	7.72 (0.47)	0.75 (0.16)	0.23 (0.06)

^aRN_CG: reaction norm with contemporary group EC, five discrete intervals for residual variance; RN_FY: reaction norm with flock-year EC, residual variance estimated for each flock-year; RN_CG_FY: reaction norm with contemporary group EC, residual variance estimated for each flock-year;

The EC used in the reaction norms was lowly correlated to the loadings of both common factors in FA2 (Figure 5.7). This indicated that the environmental variables implicit in the common factors of FA2 were very different to EC used in the reaction norm model. Therefore, we can expect the G×E interactions to be different between the models.

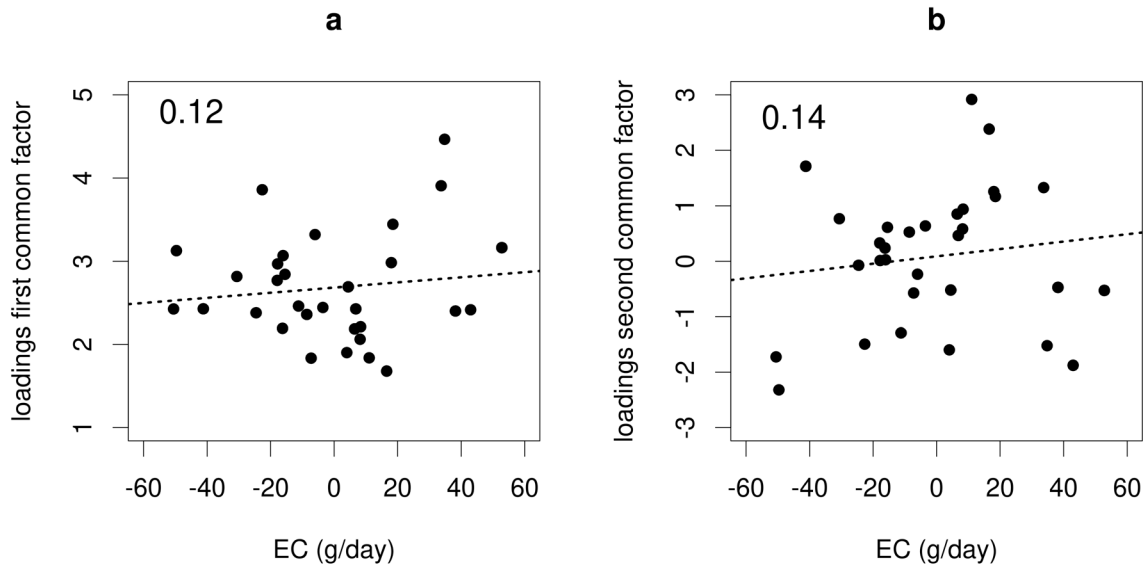


Figure 5. 7 The relationship between the average EC value within each flock-year and the loadings for the (a) first and (b) second common factor in FA2.

The genetic variance and heritability estimated for all flock-years using FA2 and RN_CG_FY are plotted in Figure 5.8. There was no relationship between the models for the estimated genetic variance of the flock-years. Compared to FA2, RN_CG_FY estimated a much smaller range of genetic variance (7.3 to 10.8 kg² vs 3.7 to 22.3 kg²) and a lower average genetic variance (8.0 kg² vs 10.2 kg²) across the flock-years. The heritability was slightly more related between models. Compared to FA2, RN_CG_FY estimated a slightly smaller range (0.24 to 0.70 vs 0.26 to 0.79) and lower average heritability (0.42 vs 0.52) across flock-years compared to FA2.

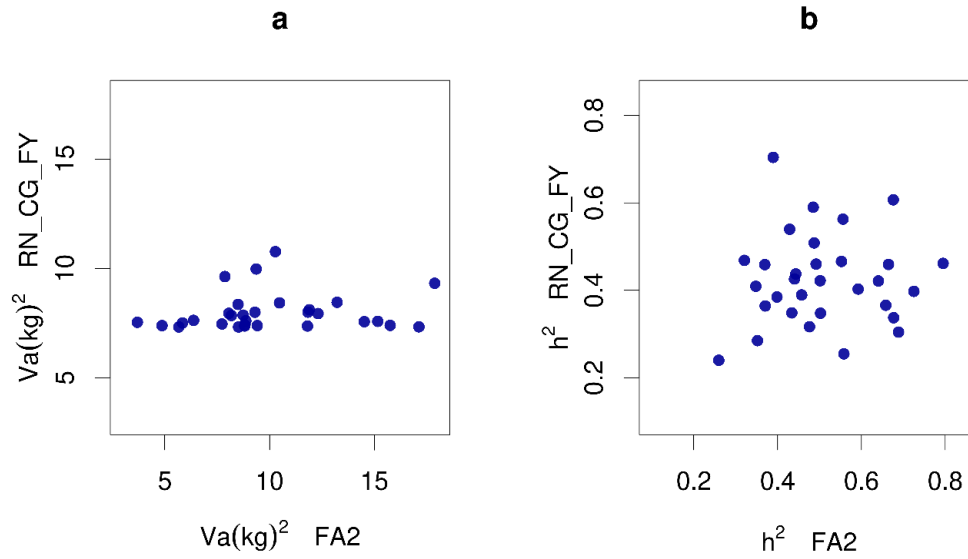


Figure 5. 8 The relationship between estimates of (a) genetic variance and (b) heritability for each flock-year, as estimated in FA2 and RN_CG_FY. The mean CG effect within each flock-year was used to obtain the estimates in RN_CG_FY.

There was no relationship between the estimates of genetic correlations amongst all of the flock-years in the data (Figure 5.9). The correlations between flocks were also compared within years, as there was more power to estimate these correlations due to higher genetic linkage. While the relationship appeared to be marginally stronger in some individual years, it was overall low. The minimum genetic correlation between flock-years was much higher in the reaction norms (0.70 vs -0.12), as well as the average correlation between all flock-years (0.96 vs 0.71). However, when considering the full range of the EC (i.e., including contemporary group variation within each flock-year), the average pairwise genetic correlation across the EC in the reaction norms was 0.85 and was as low as 0.33.

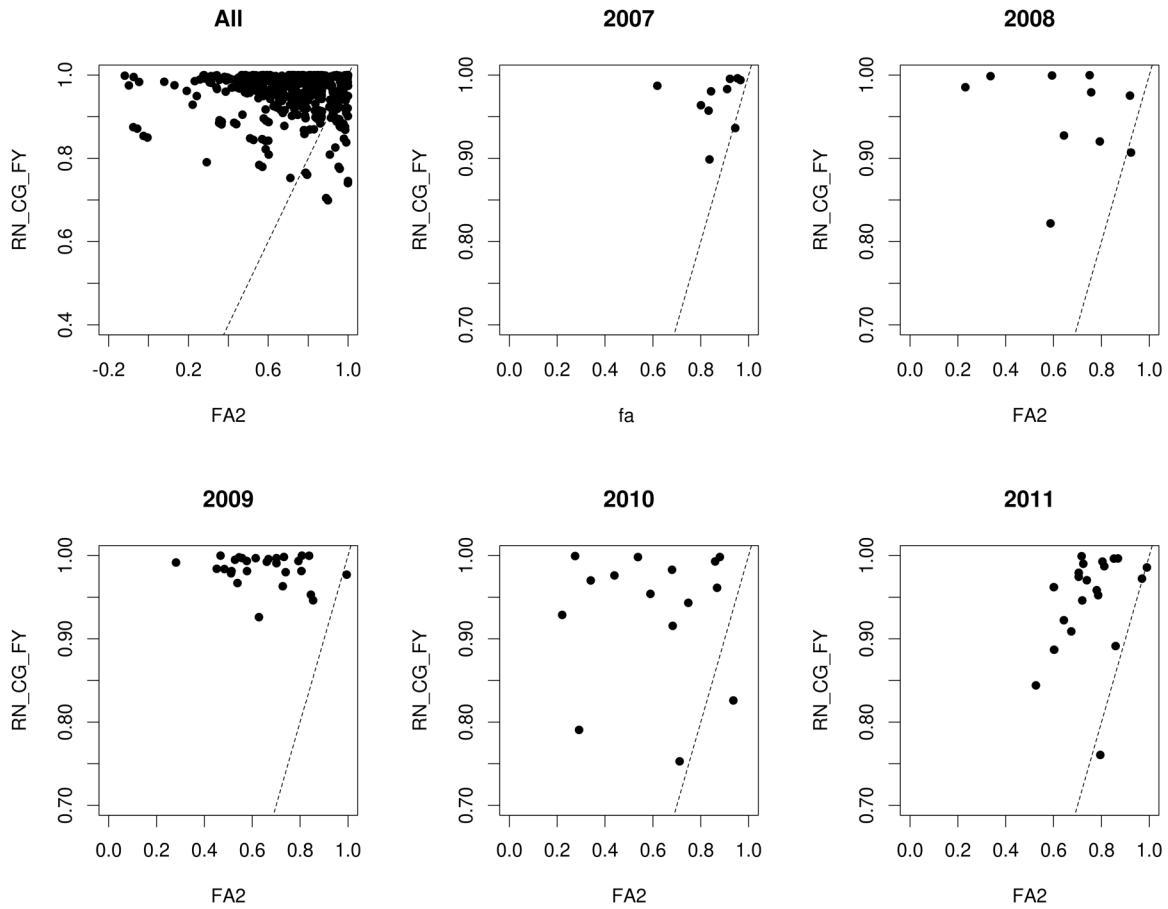


Figure 5. 9 Genetic correlations between all flock-years (top left panel) as estimated in the reaction norm model (RN_CG_FY) and factor analytic model (FA2). The correlations within years between flocks are shown separately in the remaining panels. The dashed line shows the relationship if the estimated correlations were identical between models.

The EBVs for overall performance were highly correlated (0.98) between FA2 and RN_CG_FY for a subset of 351 sires with at least 25 progeny (Figure 5.10). However, the reaction norm slope EBVs were not correlated with the responsiveness to the second common factor as estimated in the FA2 model. This was unsurprising given that the breeding values in the reaction norm model were regressed on an EC that was not related to the loadings of the second common factor in FA2 (Figure 5.7b). It was more informative to compare the absolute value of these EBVs (Figure 5.11), as this captured the robustness of each genotype across environments (i.e., the magnitude $\mathbf{R2}$, $\mathbf{a_1}$, and $\mathbf{a_1^*}$) without reference to the direction of their response across the ECs. Based on the absolute values, the rank-correlation with $\mathbf{R2}$ was small

but positive for $\mathbf{a}_1$ (0.16) and $\mathbf{a}_1^*$ (0.17). While there was a small relationship between the models, the factor analytic model tended to rank individuals differently for robustness compared to the reaction norm. Although not shown, the results were similar when comparing sires within years.

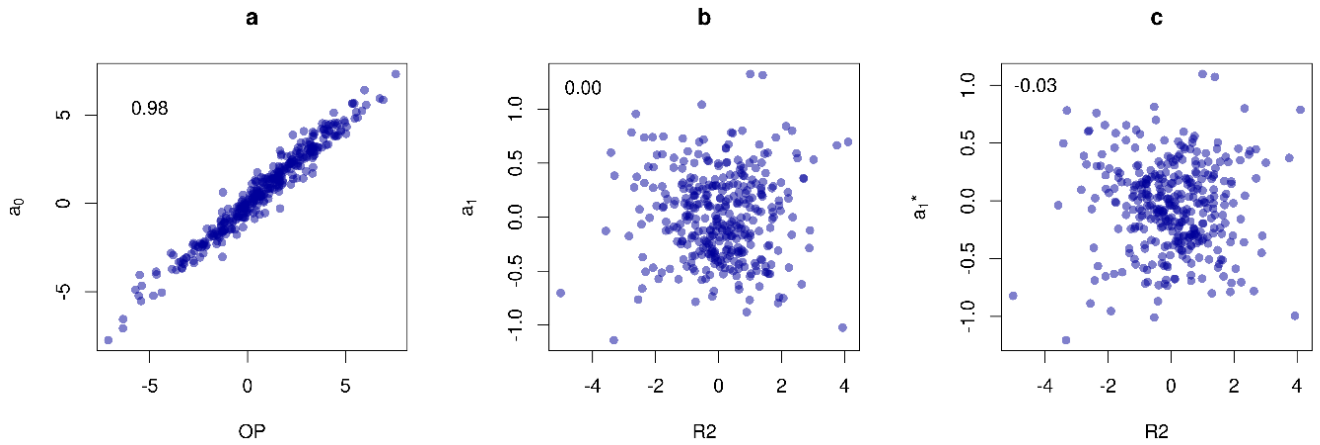


Figure 5.10 Scatterplots of EBVs for 351 sires with more than 25 progeny. The plots show (a) overall performance (**OP** vs $\mathbf{a}_0$), (b) stability using the reaction norm slope (**R2** vs $\mathbf{a}_1$) and (c) scale corrected slope (**R2** vs $\mathbf{a}_1^*$). The correlation is printed in each frame.

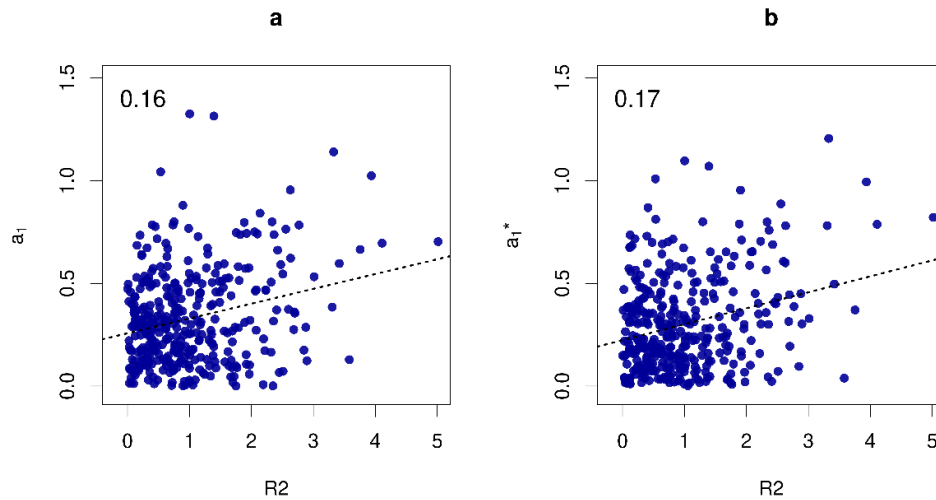


Figure 5.11 Scatterplots of the absolute value of reaction norm and factor analytic robustness EBVs for 351 sires with more than 25 progeny. The plots show (a) the reaction norm slope (**R2** vs $\mathbf{a}_1$) and (b) the scale-corrected slope (**R2** vs $\mathbf{a}_1^*$). The rank-correlation of the EBVs is printed in each frame.

5.5 Discussion

This study aimed to investigate G×E interactions in a livestock population using factor analytic models and compare the results to reaction norm models. The FA models were able to parsimoniously estimate the genetic covariance matrix between flock-years while quantifying the performance of genotypes across environmental covariables that were implicit in the data. However, the G×E interactions and breeding values for robustness as quantified by the FA models were very different to the RN models. This highlights the need for further investigation between the two different methods for analysing G×E interaction in livestock data.

In plant breeding data, the flexibility offered by several orthogonal environmental covariables to explain complex G×E interactions means that FA models often outperform competing models (e.g., Kelly et al., 2007; Smith et al., 2015; Waters et al., 2023). This was also found in the analysis of an aquaculture population (Sae-Lim et al., 2014). It is therefore conceivable that the RN estimates of G×E interactions were different from the FA models because the environmental covariable, which was the post-weaning growth of each contemporary group, did not explain much of the G×E interactions. The environmental covariable was not correlated with either of the common factors identified by the FA2 model (Figure 5.7). The reaction norm models also estimated much higher genetic correlations between environments, similar to what was observed in Waters et al. (2023). In some instances, the RN_CG_FY and FA2 model estimated a correlation of 1 and 0, respectively, between the same environments (Figure 5.9), suggesting that the reaction norm overlooked significant G×E interactions. This implies that although the environmental covariable was the same between environments, the underlying factors causing the G×E interaction were different (e.g., a lack of feed vs high parasite burden). However, it was not clear which model best fit the data based on the different model selection criteria. While the FA models were significantly better based on the AIC and log-likelihood ratio test, they were significantly worse based on the BIC, which has a more stringent criterion

for parsimony. The lack of a clear preference between the models might be a result of differences in the level of genetic linkage between environments in plant breeding and livestock data.

Out-crossing livestock populations are inherently less balanced across environments than most plant breeding populations, where identical genotypes can be assessed in different environments. Instead, the $G \times E$ interactions in livestock must be estimated based on the performance of relatives in different environments, captured by either pedigree or genomic relationships. The power of the FA models to accurately capture $G \times E$ interactions in field data of livestock is therefore highly dependent on the degree of genetic linkage between environments. In our data, genetic linkage between the flock-years was high by the standards of extensively raised livestock, as it was based on an AI program and designed with the estimation of $G \times E$ interactions in mind. Utilizing genomic relationships would have increased the connectedness between environments further, but this was computationally infeasible in this study. A study of an out-crossing pine tree population with similar levels of genetic linkage successfully modelled $G \times E$ interactions with a factor analytic model (Cullis et al., 2014). While this provides some re-assurance that the level of genetic linkage was adequate to fit the FA models, a simulation study tailored towards animal breeding data would be useful to understand the structure required to fit FA models effectively.

The FA models assumed that all animals in a flock-year were subject to the same environment. This is generally the case in plant breeding analyses because trials are stationary (i.e., they cannot change paddocks), and treatments are applied uniformly. However, this is not the case in livestock, as there are often contemporary groups within a flock-year that may experience different environmental conditions for various reasons. These potential differences were ignored in the FA models, as the contemporary groups didn't have enough individuals to be considered as an environment. This highlights a key advantage of the RN models, which could

use environmental information at the level of each contemporary group without the cost of extra parameters. The RN model which used contemporary group effects (RN CG_FY) was significantly better than the flock-year effects model (RN_FY), indicating that environmental variation within each flock-year was important. This was further underlined by examining the minimum genetic correlation between contemporary groups within each flock-year estimated in RN_CG_FY, which was as low as 0.49 and averaged 0.85. By utilising this environmental variation within flock-years, the RN models should also be less sensitive to weak genetic linkage between flock-years relative to FA models. Overall, a more definitive way to determine the utility of each model would be to test the predictive ability of their EBVs, which is the subject of future research.

The FA models provided insight into the relationship between different flock-years, which would otherwise be impossible to calculate using an unstructured matrix. The FA2 model identified two groups of environments in which performance was more highly correlated. This information is used in plant breeding programs to understand the groups of environments in which predictions of genetic merit are reliable, as well as to identify genotypes that are adapted to specific types of environments (Smith et al., 2021). It would be challenging to design a breeding program or deliver environment-specific breeding values based on the groups identified in the current study, as most of the flocks occurred within both groups. This was similar to the findings of Hardner et al. (2011), who, in a study of the diameter at breast height in Eucalyptus hybrids, discovered 3 clusters of environments that were geographically unrelated within. To be more informative, the underlying environmental factors which drive the similarity between environments need to be identified using agronomic, management, climatic and other environmental data. Alternatively, selection could be based on a global estimate of stability across all the environments.

5.6 Conclusions

Factor-analytic models can be used to explore G×E interactions and robustness in livestock data. While the estimates of overall performance were highly correlated with the reaction norm models, the G×E interactions and estimates of robustness were very different. Additional research is required to understand the appropriate data structure required to estimate G×E interactions accurately using factor analytic models in livestock.

Journal-Article Format for PhD Theses at the University of New England

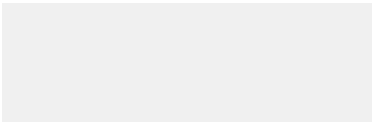
STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidates' contribution as indicated in the *Statement of Originality*.

	Author's Name	% of contribution
Candidate	Dominic Luke Waters	80
Other Authors	Sam Clark	10
	Julius van der Werf	10

Name of Candidate: Dominic Luke Waters

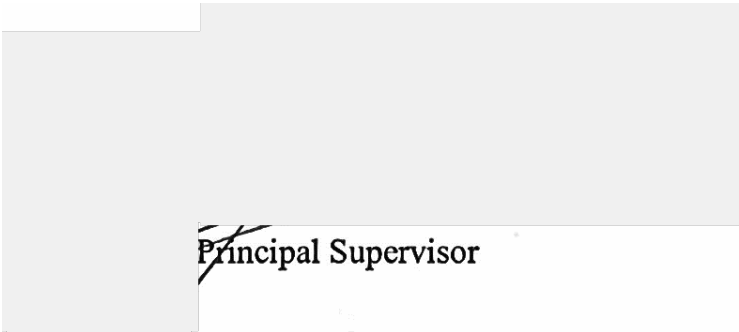
Name/title of Principal Supervisor: Professor Julius van der Werf



Candidate

30/06/2023

Date



Principal Supervisor

30/06/2023

Date

Journal-Article Format for PhD Theses at the University of New England

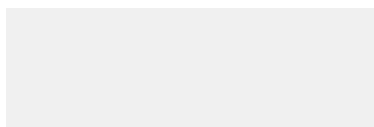
STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of Work	Page number's
Text	All
Figures	All
Tables	All

Name of Candidate: Dominic Luke Waters

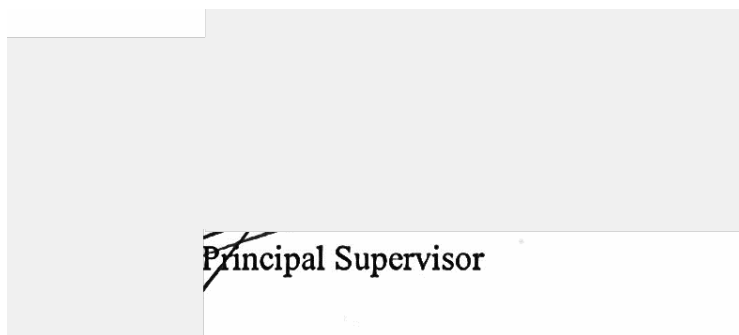
Name/title of Principal Supervisor: Professor Julius van der Werf



Candidate

30/06/2023

Date


Principal Supervisor

30/06/2023

Date

Chapter 6. Validation of reaction norm breeding values for robustness in Australian sheep

Dominic L. Waters^{1*}, Sam Clark¹, Daniel Brown², Samuel Walkom² and Julius van der Werf¹

¹School of Environmental & Rural Science, University of New England, Armidale, NSW, 2351, Australia

²Animal Genetics and Breeding Unit, University of New England, Armidale, NSW, 2351, Australia

*Corresponding author

E-mail addresses:

DLW: dwaters9@myune.edu.au

SAC: sam.clark@une.edu.au

JHVDW: jvanderw@une.edu.au

Prepared for Genetics Selection Evolution

6.1 Abstract

Background

There can be variation between animals in how stable their genetic merit is across different environments due to genotype-by-environment ($G \times E$) interactions. This variation could be used in breeding programs to select robust genotypes who combine high overall performance with stable genetic ranking across environments. There have been few attempts to validate breeding values for robustness in livestock, despite this being a necessary step towards their implementation in selection decisions. The objective of this study was to validate breeding values for the robustness of body weight across different growth environments estimated using reaction norm models in sheep data.

Results

Using three-fold cross validation for the progeny of 337 sires, the average correlation between single-step breeding values for the reaction norm slope and the realised robustness of progeny across different environments was 0.21. The correlation between breeding values for the reaction slope estimated independently in two different datasets linked by common sires was close to the expected correlation based on theory.

Conclusions

Slope EBVs estimated using reaction norm models were predictive of the phenotypic robustness of progeny across different environments and were consistent for sires with progeny in two different datasets. Selection based on reaction norm EBVs could be used to increase the robustness of a population to environmental variation.

6.2 Background

Livestock populations are often kept across a range of environments that vary due to differences in both locations and years. This can give rise to genotype-by-environment ($G \times E$) interactions, which occur when the relative performance of a genotype is dependent on the environment it will exist in. Genotype by environment interactions can be challenging to breeding programs, as it means the genetic merit of a genotype can change depending on the year or location it will be used in. However, $G \times E$ interactions also represent an opportunity, as they are a source of genetic variation from which genotypes that are more robust to environmental variation could be selected (Mulder, 2017). This genetic resource could become more valuable with climate change, as farm environments are predicted to become more variable (Rojas-Downing et al., 2017).

Genetic variation in robustness can be captured using reaction norm models, which regress the estimated breeding value (EBV) of individual genotypes across an environmental covariable (EC) that is representative of the environmental quality. A linear reaction norm estimates an intercept, which represents the average EBV across all the EC values, and a slope, which captures how much the EBV of a genotype changes over the range of the EC and is inferred from the performance of relatives in different environments. An ideal genotype combines a high intercept (high overall performance) with a flat slope (robustness) (Kolmodin et al., 2002). Variation in the slope can also be partitioned into the type of $G \times E$ interaction which underlies it; that is, heterogeneity of genetic correlation (rank-type $G \times E$), or heterogeneity of genetic variance (scale-type $G \times E$) (Waters et al., 2023). Correcting the slope EBVs for scale-type $G \times E$ could help to identify sires who have stable genetic ranking across environments.

A substantial body of research has demonstrated significant $G \times E$ interactions using reaction norm models in livestock (e.g., Cardoso & Tempelman, 2012; Carneiro et al., 2019; Nel et

al., 2023; Waters et al., 2022), as well as their potential to increase the accuracy of phenotypic predictions relative to models which ignore G×E interaction (Mota et al., 2020; Oliveira et al., 2018; Silva et al., 2014). It could also be useful to directly explore the predictive ability of reaction norm breeding values for robustness, as this would be an important step towards understanding the value of their inclusion in selection decisions.

Unlike conventional phenotypes, robustness cannot be directly measured on an individual animal. However, we can estimate the robustness of a sire's genotype by recording progeny across a range of different EC levels. By estimating sire EBVs for robustness using a training dataset, we can then examine their predictive ability in a test set. This type of cross-validation analysis has been used to validate EBVs for heat tolerance in dairy cattle (Nguyen et al., 2016), which eventually led to their implementation in selection programs (Nguyen et al., 2017). Robustness EBVs could also be validated if sires have progeny extensively recorded in two different datasets. By estimating robustness EBVs for sires separately in each dataset, the validity of robustness EBVs could be deduced by comparing the correlation of EBVs between datasets with the expected correlation based on theory.

The aim of this study was to explore the ability of linear reaction norm EBVs to predict the robustness of performance of progeny across environments using body-weight records collected on Australian Sheep. The environmental covariable was defined as the adjusted post-weaning growth rate of a contemporary group as a proxy for the quality of the environment under which growth performance was measured, and the analysed trait was post-weaning weight. Hence, robustness was captured as the change in EBV for post-weaning weight across different growth environments.

6.3 Materials and Methods

6.3.1 Data

Body weight at weaning (WWT, recorded between 50-120 days of age) and post-weaning (PWT, recorded between 120-329 days of age) was available for 34,584 lambs in the Australian Sheep CRC Information Nucleus Flock (INF) and Meat and Livestock Resource Flock (RF). The lambs were of Merino, Maternal (such as Border Leicester) or Terminal (such as Poll Dorset) sires and Merino or first cross Maternal-Merino dams. A previous genomic analysis of the reaction norm for post-weaning weight in this data revealed significant genetic variation in the reaction norm slope (Waters et al., 2022). The objective of the first part of the analysis was to test whether this genetic variation was predictive of robustness of progeny performance using cross-validation and forward prediction.

Weaning and post-weaning body weights were available on a further 344,888 pure-bred merino lambs recorded by commercial stud breeders in a national evaluation system known as MERINOSELECT (Brown et al., 2007). Many of the merino sires recorded in the INF/RF also had progeny recorded in this data. The objective of the second part of the analysis was to test whether the EBVs for sires based on the reaction norm slope in the INF/RF were consistent with the EBVs estimated in the wider industry data.

Part 1. Cross validation and forward prediction within the INF/RF dataset

As described in Waters et al. (2022) the environmental covariable (EC) was calculated for 33,773 eligible lambs as the best linear unbiased estimate (BLUE) of the post-weaning growth rate (PWGR) of each contemporary group (CG), where the CGs consisted of a flock $\times$ year $\times$ management group combination. To be eligible, each lamb required a known sire and dam, at least 40 days between weaning and post weaning weights, to be born as quadruplets and reared

as triplets or less, a member of a contemporary group with at least 15 animals originating from at least 3 different sires, and to be within 4 standard deviations of the mean PWGR. The model to obtain the BLUE of the CG effects was:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a} + \mathbf{Q}\mathbf{g} + \mathbf{e} \quad (1)$$

Where $\mathbf{y}$ is the vector of PWGR records, $\mathbf{X}$ is an incidence matrix for the fixed effects $\mathbf{b}$, $\mathbf{Z}_1$ was an incidence matrix relating the records to additive genetic effects ($\mathbf{a}$), $\mathbf{Q}$ is a matrix of the proportion of each animals genome originating from 39 breed-based genetic groups derived from their pedigree, $\mathbf{g}$ is the vector of random genetic group effects, and $\mathbf{e}$ is the vector of residual genetic effects. Fixed effects included sex, birth-type by rear-type interaction, age at post-weaning (linear and quadratic), and CG. The BLUE of the CG effects formed the EC variable in reaction norm analysis.

To be used in the validation analysis, sires required at least 25 progeny across an EC range of 60 g/day. Of the 345 sires that met this criteria, 337 were only used between 2007 and 2011, while the remaining 8 sires were used between 2012 and 2020. This was because there were more locations tested between 2007 and 2011, so sires tended to have more progeny across different EC levels. Therefore, 18,171 lambs born between 2007 and 2011 were extracted for the validation analysis. The distribution of the lambs across the EC is given in Figure 6.1.

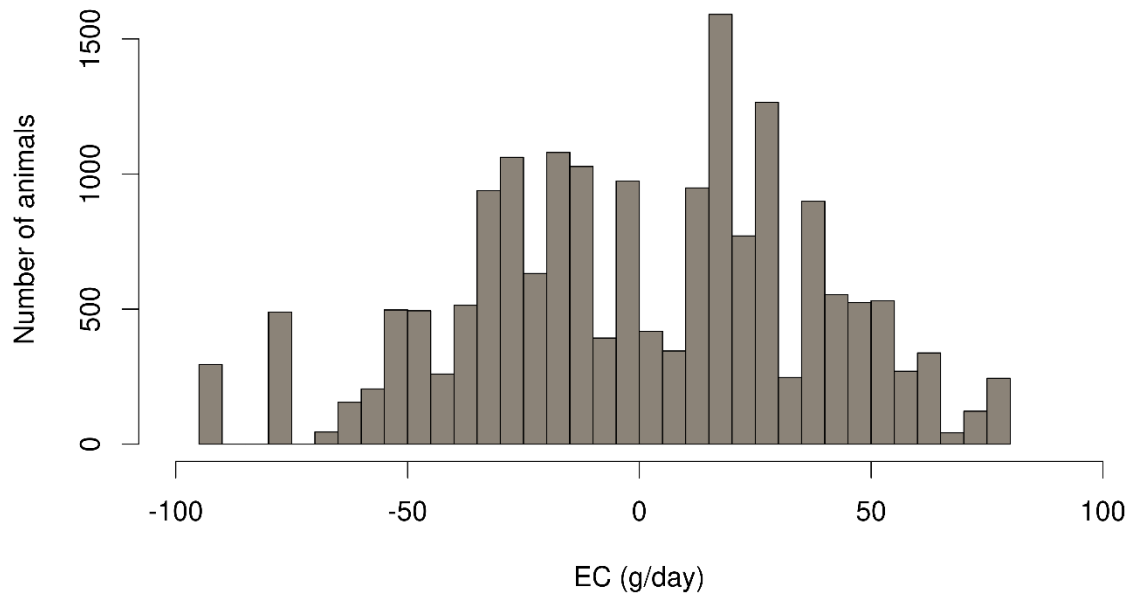


Figure 6. 1 Distribution of 18,171 lambs across the mean-centred EC.

Genomic data was available for 12,359 of the lambs, and consisted of 60,400 imputed SNPs, as described in Waters et al. (2022). Genomic information was included in the reaction norm models using the matrix **H**, which combined the pedigree (**A**) and genomic (**G**) relationship matrices (Aguilar et al., 2010). **A** was constructed using 3 generations of pedigree (31,502 animals), while **G** was constructed following the first method proposed in VanRaden (2008). The two matrices were combined to form **H** using MTG2 (Lee & van der Werf, 2016).

6.3.2 Cross-validation

In the three-fold cross-validation, the progeny of each sire was split into three folds. To ensure each fold contained the maximum environmental range possible, progeny were ordered by the EC within each sire. The first three progeny were then randomly assigned to a fold (1, 2 or 3), without replacement. This was repeated for the next three progeny along the EC, and so on for the remainder of the progeny. Only sires that had an environmental range of at least 60 g/day within each of the three folds were considered in the cross-validation, leaving 337 sires of 14,612 lambs.

The single-step reaction norm models were trained on progeny from two of the folds, as well as the progeny of sires who did not meet the criteria for cross-validation (3559 lambs from 142 sires). The resulting GEBVs were then used to predict the realised progeny performance in the remaining fold for the 337 validation sires. This was repeated for each of the three folds. A numerical summary for the three folds given in Table 6.1, as well as the additional data from the progeny of ineligible sires used to train the models.

Table 6. 1 Numerical summary of animals within each fold, as well as the progeny of sires not eligible for validation but used in the training data (Additional training). The mean $\pm$ standard deviation is reported for post-weaning weight and the EC.

Fold	1	2	3	Additional training
No. Animals	4876	4880	4856	3559
No. genotyped animals	3331	3338	3294	2396
Post-weaning weight (kg)	43.1 $\pm$ 8.7	43.1 $\pm$ 8.8	43.1 $\pm$ 8.6	41.0 $\pm$ 9.1
EC (g/day)	0.86 $\pm$ 38.2	0.68 $\pm$ 38.2	0.76 $\pm$ 38.1	-3.14 $\pm$ 32.9

The reaction norm model fitted was:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a}_0 + \mathbf{Z}_2\mathbf{a}_1 + \mathbf{Z}_3\mathbf{c} + \mathbf{Q}\mathbf{g} + \mathbf{e} \quad (2)$$

Where $\mathbf{y}$ is a vector of PWT phenotypes, $\mathbf{X}$ is an incidence matrix for the fixed effects $\mathbf{b}$, $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are matrices relating records to the additive genetic effects for the intercept ($\mathbf{a}_0$) and slope ($\mathbf{a}_1$) respectively, with $\mathbf{Z}_1$ containing 1's on the diagonal and $\mathbf{Z}_2$ containing the EC value corresponding to each individual in $\mathbf{y}$ on the diagonal, $\mathbf{Z}_3$ is an incidence matrix relating records to the dam effects ($\mathbf{c}$), which were estimated ignoring relationships between dams, $\mathbf{Q}$ is a matrix of the proportion of each animals genome originating from 39 breed-based genetic groups, $\mathbf{g}$ is the vector of random genetic group effects, and $\mathbf{e}$ is the vector of residual genetic effects. Fixed effects included age at measurement (linear and quadratic), birth-type and rear-type

interaction, sex, and contemporary group. The additive genetic variance of $\mathbf{a}_0$ and $\mathbf{a}_1$ was modelled according to $\begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \end{bmatrix} \sim N(0, \mathbf{H} \otimes \mathbf{K})$, where $\mathbf{K} = \begin{bmatrix} \sigma_{a_0}^2 & \sigma_{a_1 a_0} \\ \sigma_{a_0 a_1} & \sigma_{a_1}^2 \end{bmatrix}$. The residual ($\mathbf{e}$) was modelled as a continuous function of the EC rather than using discrete intervals. An intercept ($\mathbf{e}_1$) and slope ($\mathbf{e}_2$) residual coefficient was used, such that $\begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \end{bmatrix} \sim N(0, \mathbf{I} \otimes \mathbf{E})$, where $\mathbf{E} = \begin{bmatrix} \sigma_{e_1}^2 & \sigma_{e_2 e_1} \\ \sigma_{e_1 e_2} & \sigma_{e_2}^2 \end{bmatrix}$. The models were implemented using MTG2 (Lee & van der Werf, 2016).

6.3.3 Forward prediction

In the forward prediction analysis, data from 14,521 lambs born in 2007 to 2010 were used to train the reaction norm model and predict the progeny performance of 88 sires with 3103 progeny born in 2011. Unlike the cross-validation, the progeny of validation sires were not used to train the reaction norm model, so they had a comparably distant relationship to the training population. The same reaction norm model as Eq. 2 was fit to the data. The distribution of the training and test data across the EC is given in Figure 6.2.

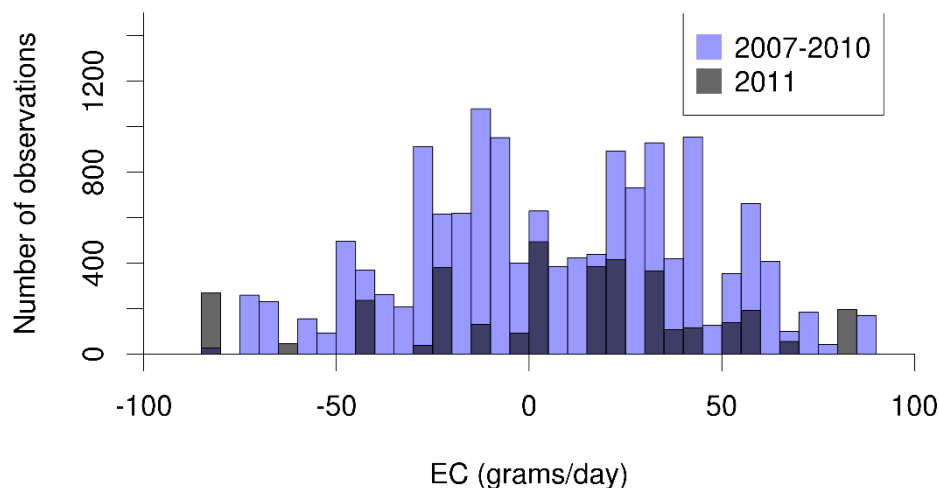


Figure 6. 2 Distribution of lambs in 2007-2010 (blue) and 2011 (grey) across the mean-centred EC.

6.3.4 Realised performance

The realised progeny performance across the EC was obtained as a linear random regression of the pre-adjusted phenotypes ($\mathbf{y}^*$) of the sire's test progeny across the EC. The pre-adjustment factors were calculated using an animal model:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{c} + \mathbf{Qg} + \mathbf{e} \quad (3)$$

Where $\mathbf{y}$ is the vector of PWT records, $\mathbf{X}$ is an incidence matrix for the fixed effects $\mathbf{b}$, which were the same as in Eq. 2, $\mathbf{Z}_1$ was an incidence matrix relating the records to additive genetic effects ($\mathbf{a}$), and $\mathbf{Z}_2$, $\mathbf{Q}$, $\mathbf{c}$, $\mathbf{g}$ and $\mathbf{e}$ were the same as in Eq.1. The pre-corrected phenotypes ($\mathbf{y}^*$) were then obtained as:

$$\mathbf{y}^* = \mathbf{y} - \mathbf{Xb} - \mathbf{Z}_2\mathbf{c} - \mathbf{Qg} \quad (4)$$

The pre-corrected phenotypes were regressed across the EC for each sire using a linear random regression model. The model ignored the relationships between sires and estimated the residual variance as heterogenous at three discrete intervals along the EC (EC = -91 to -26 g/day, -26 to +26 g/day and 26 to 79 g/day). The resulting estimates of the intercept ($\mathbf{p}_0$) and slope ($\mathbf{p}_1$) captured the overall phenotypic performance and robustness of each sire's progeny across the EC, respectively. The Pearson product-moment correlation between the GEBVs for sires in the training set and the realised progeny performance estimated from the test set was used to quantify the predictive ability of the GEBVs.

Part 2. Validation of robustness EBVs in industry data

Many of the Merino sires used in the INF/RF also had progeny recorded in the wider industry population, known as MERINOSELECT (Brown et al., 2007). The aim of this section was to test whether reaction norm slope EBVs estimated in the INF/RF data were consistent with equivalent EBVs estimated in the MERINOSELECT data. The analysis used pedigree data to

model additive relationships because of both computational limitations and potential issues with selective genotyping in the commercial data.

6.3.5 Data

There were 318,028 pure-bred Merino lambs in the MERINOSELECT database that were eligible to estimate EC. Additionally, CGs in the MERINOSELECT data were only included if they contained at least one animal related to a sire or grandsire in the INF/RF data, leaving 277,060 lambs. There were 12,087 pure-bred Merino lambs eligible to estimate the EC in the INF/RF data. The same model as Eq. 1 was used to obtain a BLUE of PWGR for each CG using the joint MERINOSELECT and INF/RF datasets, except that an additional dam effect (c) was fitted.

After mean-centring, the EC effects ranged between -200.9 and +256.5 /day in MERINOSELECT and -99.2 and +122.7 g/day in the INF/RF (Figure 5.3). Only CGs within ± 120 g/day (MERINOSELECT), and ± 75 g/day (INF/RF) of the mean were used in the reaction norm analysis, because small amounts of data at the ends of an EC can produce unreliable estimates in reaction norm models. To further reduce the number of uninformative animals in the MERINOSELECT data, CGs were only included if they contained 1) at least one progeny of a sire with progeny or grand-progeny in the INF/RF, or 2) more than 25% of the animals related to the INF/RF, with the minimum relationship being a grandsire with grand-progeny in the INF/RF. This left 11,638 and 198,331 animals for the reaction norm analysis in the INF/RF and MERINOSELECT datasets, respectively. There were 253 common sires with 23,150 and 4729 direct progeny in the MERINOSELECT and INF/RF datasets, respectively. In total, 114,918 animals in MERINOSELECT and 5291 animals in the INF/RF were at least great-grand progeny of the common sires (i.e., shared at least ~12.5% of their genome as identical by descent). A numerical summary of the final INF/RF and MERINOSELECT data is given in Table 6.2.

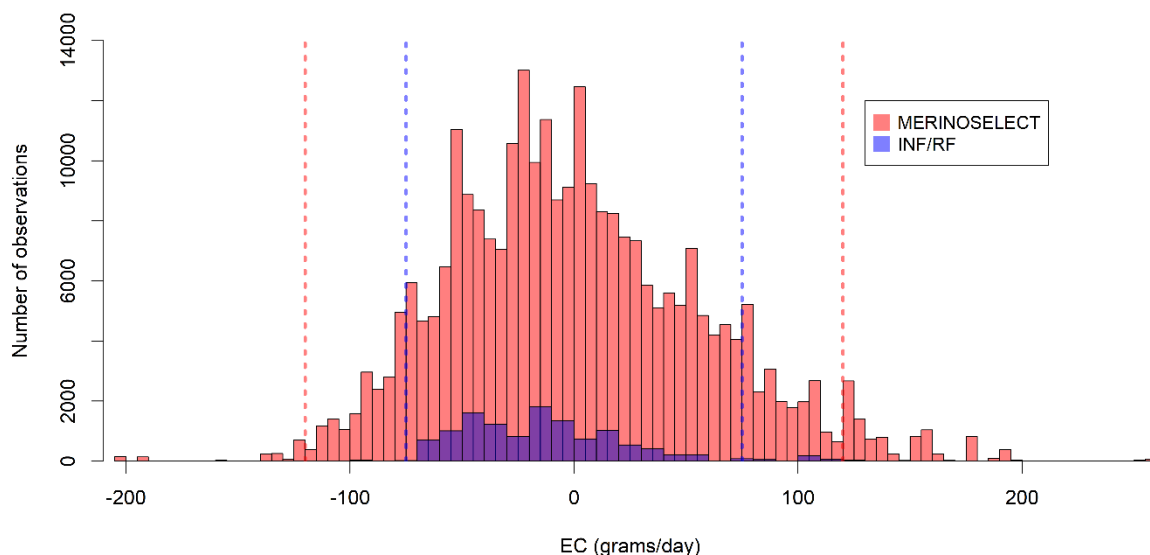


Figure 6. 3 Distribution of lambs across the EC as a deviation from the overall mean EC in the MERINOSELECT (red) and INF/RF (purple) datasets. The vertical lines show the range of data used to fit the reaction norm models in the two datasets.

Table 6. 2 Numerical summary of animals in the INF/RF and MERINOSELECT

	INF/RF	MERINOSELECT
Number of lambs	11,638	198,331
Number of sires with progeny data	615	5389
Number of dams	7070	116978
Number of CG	134	1791
Weaning age (days)	94 ± 9.8	94.8 ± 14.7
Post-weaning age (days)	262.0 ± 35.5	228.3 ± 40.3
Post-weaning growth period (days)	167.9 ± 36.9	133.5 ± 40.9
Weaning weight (kg)	24.5 ± 5.3	27.7 ± 6.7
Post-weaning weight (kg)	38.4 ± 8.05	40.3 ± 10.1
Post-weaning growth rate (g/day)	84.1 ± 39.4	96.7 ± 57.5
EC (g/day)	-16.5 ± 29.8	-4.4 ± 48.5

6.3.6 Reaction norms

Linear reaction norms were fit to the INF/RF and MERINOSELECT datasets independently to model how EBVs for PWT changed across the EC using ASReml 4.2 (Gilmour et al., 2021).

The linear reaction norm models were of the form:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a}_0 + \mathbf{Z}_2\mathbf{a}_1 + \mathbf{Z}_3\mathbf{c} + \mathbf{Z}_4\mathbf{m} + \mathbf{Q}_g + \mathbf{e} \quad (5)$$

Where $\mathbf{y}$, $\mathbf{X}$, $\mathbf{Z}_1$, $\mathbf{Z}_2$, $\mathbf{Z}_3$, $\mathbf{Q}$, $\mathbf{b}$, $\mathbf{a}_0$, $\mathbf{a}_1$, $\mathbf{c}$, $\mathbf{g}$ and $\mathbf{e}$ were the same as in Eq. 2, and $\mathbf{Z}_4$ is an incidence matrix relating records to the additive maternal effects ($\mathbf{m}$). The additive maternal genetic effects were not fitted in the INF/RF model as they were not significant. There were 451 different genetic groups which were formed on a flock and time basis (see Brown et al., 2007).

ASReml 4.2 does not allow the residual variance to be modelled as a continuous function of the EC, so it was estimated at four and six discrete classes along the EC for the INF/RF and MERINOSELECT dataset, respectively. The variance of the intercept and slope was modelled

as follows: $\begin{bmatrix} \mathbf{a}_0 \\ \mathbf{a}_1 \end{bmatrix} \sim N(0, \mathbf{A} \otimes \mathbf{K})$ where $\mathbf{K} = \begin{bmatrix} \sigma_{a0}^2 & \sigma_{a1a0} \\ \sigma_{a0a1} & \sigma_{a1}^2 \end{bmatrix}$ and $\mathbf{A}$ is the pedigree relationship matrix.

To compare estimates of genetic variance and heritability between the datasets along the different levels of the EC, the genetic (co)variance between breeding values at different levels of the EC was obtained using $\mathbf{G} = \mathbf{AKA}'$, where $\mathbf{A}$ contained two columns; the first was a vector of 1's, and the second was a vector of EC values. The heritability of PWT at a given EC level was then obtained by dividing the genetic variance at the EC level by the sum of the genetic, maternal genetic (if fitted), permanent environmental dam, and residual variance at the EC level. Scale-corrected EBVs for the slope ($\mathbf{a}_1^*$) were estimated within each dataset using a

genetic regression, which captures variation in the slope that is independent of the genetic correlation between the intercept and slope:

$$\mathbf{a}_1^* = \mathbf{a}_1 - \frac{\sigma_{a_0 a_1}}{\sigma_{a_0}^2} \mathbf{a}_0 \quad (6)$$

Sire EBVs for the intercept ($\mathbf{a}_0$), slope ($\mathbf{a}_1$) and scale corrected slope ($\mathbf{a}_1^*$) were compared between the two datasets using the Pearson product-moment correlation. This was compared to the expected correlation, which is calculated as the product of the average EBV accuracies of the two datasets. The accuracy of each reaction norm EBV was calculated as $\mathbf{r} = \sqrt{1 - \frac{\text{PEV}}{\sigma_a^2}}$, where PEV is the prediction error variance of the EBV, and σ_a^2 is the estimated genetic variance associated with either $\mathbf{a}_0$ or $\mathbf{a}_1$. The average accuracy of the reaction norm EBVs for each dataset was obtained using Fischer's Z transformation. The expected correlation was calculated only for the intercept ($\mathbf{a}_0$) and slope ($\mathbf{a}_1$) EBVs that were not corrected for scale, as further investigation is required to derive the **PEV** for $\mathbf{a}_1^*$. Because $\mathbf{a}_1^*$ was derived from $\mathbf{a}_1$, the expected correlation for $\mathbf{a}_1$ served as an approximate expectation for $\mathbf{a}_1^*$.

6.6 Results

Part 1. Cross validation and forward prediction within the INF/RF dataset

In the cross-validation analysis, the average correlation between EBVs for the intercept and the realised overall progeny performance was high (0.66, Table 6.3). The EBVs for the reaction norm slope and scale-corrected slope were moderately correlated with the realised robustness of progeny performance across the EC (0.19-0.21). Using forward prediction, the correlation between the reaction norm slope and the realised progeny robustness was similar to the cross-validation analysis (0.23) but was comparably lower for the intercept (0.53) and scale-corrected slope (0.10). Overall, the results show that the single-step reaction norm EBVs for the intercept, slope, and scale-corrected slope were predictive of the robustness of progeny performance across the EC.

The variance components estimated in each of the models are also reported in Table 6.4. Notably, the correlation between the intercept and slope was small and not significantly different from zero in all models. This suggests that there were only small scale-type G×E interactions in the data, so the results for the scale-corrected slope ($\mathbf{a}_1^*$) should be interpreted cautiously.

Table 6. 3 Correlation between reaction norm EBVs for sires ($\mathbf{a}_0$, $\mathbf{a}_1$ and $\mathbf{a}_1^*$), and the realised intercept ($\mathbf{p}_0$) and slope ($\mathbf{p}_1$) of the random regression of pre-corrected progeny phenotypes across the EC in the test sets (3-folds). The results for the forward prediction are also reported.

Fold	$\mathbf{a}_0 \text{ v } \mathbf{p}_0$	$\mathbf{a}_1 \text{ v } \mathbf{p}_1$	$\mathbf{a}_1^* \text{ v } \mathbf{p}_1$
1	0.70	0.14	0.13
2	0.63	0.23	0.17
3	0.68	0.26	0.27
Average	0.67	0.21	0.19
Forward prediction	0.53	0.23	0.10

Table 6. 4 Variance of additive genetic intercept (σ_{a1}^2), slope (σ_{a2}^2) and the correlation between them (r_{a1a2}), as well as genetic group (g) and permanent environmental dam (c) variance for the reaction norms models fitted to each fold in the cross-validation, the forward prediction (2007-2010), and all of the data (18 171, lambs). The standard errors are given in brackets.

Data	σ_{a1}^2	σ_{a2}^2	r_{a1a2}	g	c
Fold 1	8.12 (0.51)	0.65 (0.23)	0.08 (0.11)	40.9 (14.20)	2.07 (0.28)
Fold 2	7.06 (0.48)	0.81 (0.24)	0.07 (0.10)	40.80 (14.23)	2.35 (0.28)
Fold 3	7.85 (0.51)	0.66 (0.23)	-0.01 (0.11)	45.12 (15.90)	2.05 (0.28)
2007 - 2010	7.64 (0.47)	0.67 (0.21)	0.08 (0.10)	39.26 (13.70)	2.27 (0.25)
All data	7.59 (0.32)	0.63 (0.18)	0.04 (0.09)	41.13 (14.35)	2.20 (0.22)

Part 2. Validation in industry data

The variance components estimated using linear reaction norms for pure-bred Merinos in the INF/RF and MERINOSELECT datasets are reported in Table 6.5. The INF/RF had greater additive genetic variation in both the intercept and slope, while the ratio of slope to intercept variance was also greater in the INF/RF, meaning that the G×E interactions were relatively larger than in MERINOSELECT. The genetic correlation between the intercept and slope was very similar, as well as the permanent environmental dam effects.

Table 6. 5 Variance of the additive genetic intercept (σ_{a1}^2), slope (σ_{a2}^2), the ratio ($\sigma_{a2}^2/\sigma_{a1}^2$) and the correlation between them (r_{a1a2}), as well as genetic group (g), permanent environmental dam (c), and maternal genetic (m) variance for the reaction norms estimated in the INF/RF and MERINOSELECT (MS) data.

Model	σ_{a0}^2	σ_{a1}^2	$\sigma_{a1}^2/\sigma_{a0}^2$	r_{a1a2}	g	c	m
INF/RF	11.35	1.98	0.17	0.64	7.6	1.64	-
MS	8.11	1.01	0.12	0.62	3.96	1.24	1.08

The genetic variance and heritability were consistently higher across the EC in the INF/RF data (Figure 6.4). However, the pattern of increase along the EC was very similar, indicating large scale-type G×E interactions.

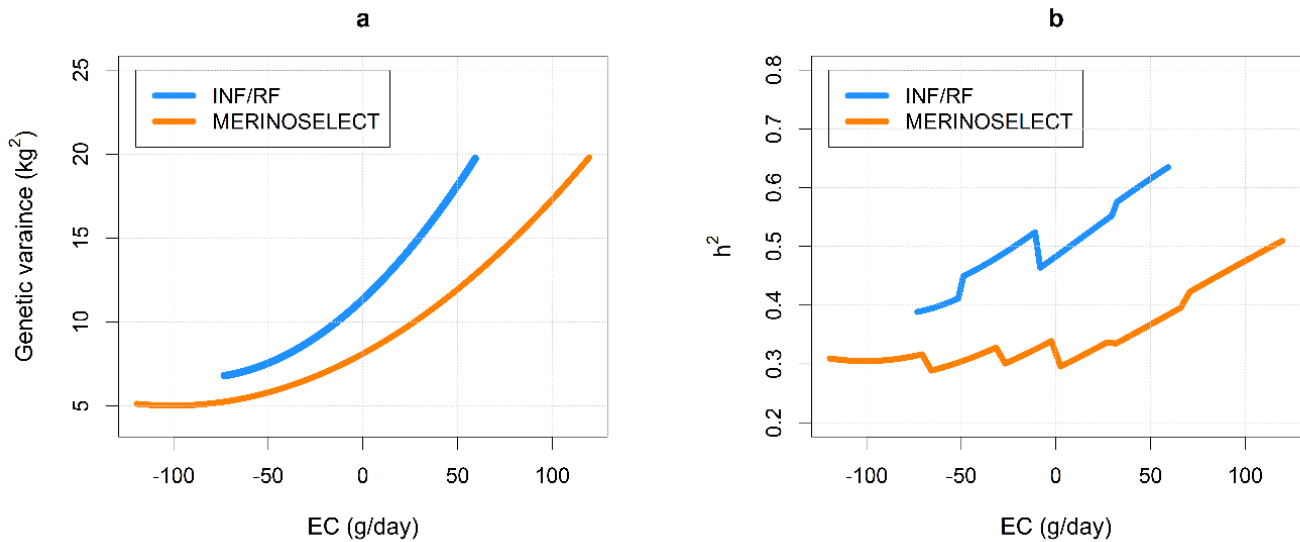


Figure 6. 4 Genetic variance (a) and heritability (b) across the EC in the MERINOSELECT and INF/RF datasets.

Using all sires with progeny in both datasets, the correlation for the reaction norm slope and scale-corrected slope EBVs were 0.24 and 0.14, respectively (Table 6.6), while the correlation between intercept EBVs was higher (0.43). The correlations for all EBVs increased when only considering sires with either the most progeny or the most contemporary groups.

The correlations for the slope were only slightly lower than those expected based on quantitative genetic theory. Since the expected correlation represents the upper bound given the accuracy of the EBVs, we can conclude that the EBVs for robustness were consistent across the two independent datasets.

Table 6. 6 Correlation of reaction norm EBVs for intercept ($\mathbf{a_0}$), slope ($\mathbf{a_1}$) and scale-corrected slope ($\mathbf{a_1^*}$) for sires with (a) direct progeny in both datasets, (b) at least 45 progeny in the MERINOSELECT and 15 progeny in the INF/RF, (c) progeny in at least 4 CGs in both datasets. The expectation of the correlation based on quantitative genetic theory is given in brackets.

EBV	a	b	c
$\mathbf{a_0}$	0.43 (0.67)	0.64 (0.75)	0.63 (0.77)
$\mathbf{a_1}$	0.24 (0.31)	0.33 (0.36)	0.31 (0.37)
$\mathbf{a_1^*}$	0.14	0.31	0.31
No. sires	253	65	62

6.7 Discussion

This study aimed to validate breeding values for robustness estimated using reaction norm models in Australian sheep. The results showed that reaction norm EBVs for the slope, which can be interpreted as ‘robustness EBVs’, were predictive of the robustness of phenotypes of progeny across a range of environments. Additionally, the reaction norm EBVs were consistent across different datasets. Therefore, reaction norm EBVs could be used to select sires who yield lambs that have weight gain that is more robust to environmental variation.

The closest comparable validation analysis of robustness to our study examined heat tolerance in dairy cattle (Nguyen et al., 2016). In their analysis, the robustness of milk, fat and protein yields to a temperature-humidity index was calculated for each cow using a linear random regression model to estimate a slope while ignoring genetic relationships. The robustness of sires was then obtained as the combination of their daughters’ slopes, and heat tolerance GEBVs were estimated using the sire slopes as phenotypes in a GBLUP model. Using forward-prediction with a reference population of 2300 Jersey sires, the correlation between heat tolerance GEBVs and the realised heat tolerance ranged between 0.28 and 0.31 for the first parity. The forward prediction strategy in our study was close to their validation structure and

yielded a slightly lower correlation of 0.23, which is probably because our study had substantially fewer sires (337 sires vs 2300 sires) and measured animals (18,171 lambs vs 366,835 cows). From this perspective, the results of the current study are promising, given that heat-stress GEBVs are now successfully applied in dairy breeding programs (Nguyen et al., 2017).

To evaluate the correlations obtained in the cross-validation analysis, 337 sires with 43 progeny each were simulated across an EC, assuming the same variance components as estimated for all the data (Table 6.4). The simulation assumed that sires were unrelated and were randomly mated to dams distributed across the EC. Across 20 runs with 3-fold cross validation, the average correlation between training and test sets was 0.66 for the intercept, 0.14 for the slope, and 0.13 for the scale-corrected slope using random regression. This was consistent with the results presented in Table 6.3, although the empirical correlation was slightly higher for the reaction norm slope EBVs. This was probably because we modelled the relationships between animals using genomic information, while the simulation assumed that sires were unrelated. This gives further reassurance that the results in the current study demonstrate the potential of selecting reaction norm EBVs to improve robustness.

The correlations between reaction norm EBVs estimated in the INF/RF and MERINOSELECT datasets were slightly lower than expected by theory. However, the theoretical correlation assumed that the genetic correlation between the datasets was one. The small deviations from theory were consistent with a genetic correlation between datasets of approximately 0.70 ~ 0.80 for the intercept, and 0.85 for the slope. This could be reasonable given that the MERINOSELECT models were trained on lambs in a much wider range of EC values (i.e., different environments), which could result in G×E interactions occurring between the datasets. Therefore, the analysis demonstrated that reaction norm EBVs for robustness in PWT to

different growth environments were consistent across two independent datasets and could be used for selection.

The correlation of reaction norm EBVs between the INF/RF and MERINOSELECT datasets was higher for the top 25% of sires with the best distribution of progeny in both datasets. This highlighted the importance of data structure to accurately capture genetic variation in robustness. Sires require progeny across a wide range of EC values to estimate the reaction norm slope accurately (Calus et al., 2004). Therefore, the success of breeding for robustness will likely be dependent on the structure of data available to estimate it accurately. If robustness is to be considered in genetic evaluations, breeders should be encouraged to ensure even stronger genetic linkage across years and locations. Alternatively, GEBVs for robustness could be supplied using an appropriately designed reference population, such as the INF/RF used in this study. Selection based on GEBVs for traits only measured within the INF/RF (such as intramuscular fat) has yielded significant responses in the national flock (Sheep Genetics, 2022), so a similar response could conceivably be gained for robustness. However, the INF/RF data had a much narrower range of EC values compared to the MERINOSELECT data (Figure 1), which could limit its ability to predict robustness in PWT to different growth environments.

Another way of estimating breeding values for robustness would be to consider the EC used to regress EBVs on. While we used the adjusted post-weaning growth rate of the contemporary groups, other ECs could be derived from measured environmental variables. This approach appears to work well for traits that have a short time in which a small number of environmental factors are important in influencing the phenotype. For instance, Nel et al. (2023) found significant changes in the genetic variation of lamb survival as a function of a cold stress EC, which was calculated as a function of temperature, wind, and rainfall. Acute cold stress was an ideal EC to regress on, as it is a critical factor influencing neonatal lamb survival (Donnelly et al., 1985; Obst & Day, 1968; Oldham et al., 2011). The regression of milk production on a

temperature-humidity index in dairy cattle is also ideal for these reasons (West, 2003). In contrast, the post-weaning weight of a lamb is affected by many interacting environmental factors over a six-month period, which makes it hard to identify and parameterise the relevant environmental descriptors. However, the adjusted post-weaning growth rate of a contemporary group automatically indexes the environmental factors into a single value with a practical interpretation; it captures the unit change in genetic merit of an individual for every unit change in growth environment. It also does not require any additional investment to capture since it is derived from phenotypes routinely collected in breeding programs. The results of the study demonstrated that the differences in robustness of sire EBVs to a post-weaning growth EC were repeatable, meaning that it captured meaningful G×E interactions and could be used for selection.

The scale-corrected slope EBVs represented slope variation that was independent of the genetic correlation with the intercept. This meant they captured slope variation that was available for selection independent of the overall performance. In the across-dataset analysis, the genetic correlation between intercept and slope was reasonably large (0.64-0.62), which meant that the scale correction (Eq. 6) had a large effect on the slope EBVs. When considering the sires with the best distribution of progeny across the EC, the scale corrected EBVs were as highly correlated between datasets as the raw slope values. This indicated there was significant genetic variation in robustness that could be selected alongside overall performance in PWT.

In the validation within the INF/RF data, the scale-corrected slope did not improve the predictive ability relative to the uncorrected reaction norm slope. However, the genetic correlation between the intercept and slope across all the data was 0.07, implying that only 0.49% ($0.07^2 \times 100$) of the genetic variation in the slope was due to scale-type G×E (Falconer, 1990). The scale-correction would have limited value in this situation, so the results should be treated with caution. The smaller correlation could indicate that modelling the residual as a

continuous function was able to adjust for the changes in phenotypic variance better than the discrete classes used in the across-dataset analysis, as suggested by (Lillehammer et al., 2009).

6.8 Conclusions

Differences between sires in their progeny performance across environments can be predicted using reaction norm models. Selection based on EBVs from reaction norm models could be used to improve the robustness of weight gain in Australian sheep to different growth environments. The success of selection for robustness will depend on the availability of appropriate data structures to estimate it. If selection for robustness is to be implemented, sires need to be used across a wide range of environments to accurately estimate their EBVs.

Journal-Article Format for PhD Theses at the University of New England

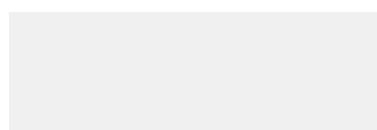
STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidates' contribution as indicated in the *Statement of Originality*.

	Author's Name	% of contribution
Candidate	Dominic Luke Waters	78
Other Authors	Sam Clark	8
	Daniel Brown	4
	Samuel Walkom	2
	Julius van der Werf	8

Name of Candidate: Dominic Luke Waters

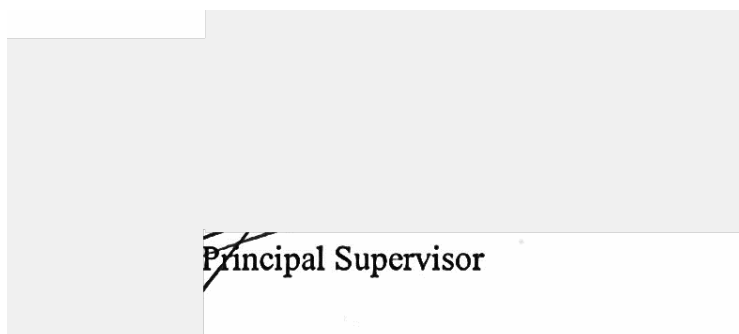
Name/title of Principal Supervisor: Professor Julius van der Werf



Candidate

30/06/2023

Date



Principal Supervisor

30/06/2023

Date

Journal-Article Format for PhD Theses at the University of New England

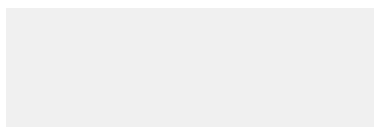
STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of Work	Page number's
Text	All
Figures	All
Tables	All

Name of Candidate: Dominic Luke Waters

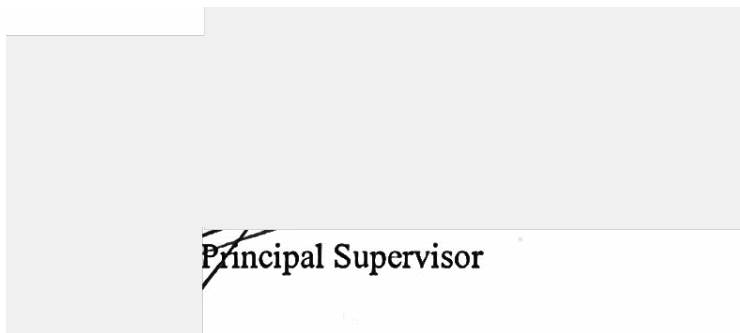
Name/title of Principal Supervisor: Professor Julius van der Werf



Candidate

30/06/2023

Date


Principal Supervisor

30/06/2023

Date

Chapter 7. General Discussion

This thesis aimed to estimate of genotype by environment interactions in livestock and plant populations using different methods and improve our understanding of how these interactions could be used in breeding programs to increase the robustness of agricultural populations to environmental variation. There has been a long-standing interest in the selection of genotypes that have performance that is more robust to environmental variation in both animals (Falconer, 1990; James, 1961) and plants (Finlay & Wilkinson, 1963; Perkins & Jinks, 1968; Yates & Cochran, 1938a). Although the robustness of genotypes to various environmental factors is often considered in the selection of plant cultivars (Bapela et al., 2022; Khanna et al., 2022; Mahjourimajd et al., 2016; Mart et al., 2016; Telfer et al., 2022), it has not been applied in animal breeding programs in the same way (Berghof et al., 2019). This is partly due to the relative ease in which identical copies of inbred plants could be grown in different environments. Without the need for computationally demanding pedigree or genomic based relationship matrices, G×E interactions (and robustness) of inbred plants could be studied in detail from a much earlier stage than in livestock (Kempton, 1984; Knight, 1970; Nabugoomu et al., 1999). This hurdle is, however, becoming less significant over time. The increasing size and genetic connectivity of livestock datasets across different environments combined with the computational resources to analyse them has led to the detection of significant G×E interactions across a variety of livestock species (Hayes et al., 2016; Rauw & Gomez-Raya, 2015). However, more research is required to quantify the magnitude of G×E interactions for important traits in livestock species and to improve our understanding of how different models can capture genetic variation in the robustness of genotype performance across environments. This may allow us to better interpret the underlying mechanisms of G×E interactions and enable breeding programs to have more confidence to implement selection for robustness.

The third chapter of this thesis investigated genotype by environment interactions and genetic variation in robustness of Australian sheep using linear reaction norms and multi-trait models in combination with genomic data. The analysis showed there was significant genetic variation in the slope of the reaction norm for post-weaning weight across different growth environments, suggesting this could be used to select for sheep that have genetic merit for post-weaning weight that is more robust to different growth environments. The reaction norm slope cannot be observed from single phenotypes, so traditional genome-wide association studies where genotypes are associated with phenotypes are not possible. However, a way around this is to back-solve SNP effects from the genomic breeding values for the reaction norm slope.

A cluster of SNPs that were significantly associated with the reaction norm slope was identified on chromosome 7. These SNPs had an increased effect in low-growth environments and occurred within the *SLC24A1* gene, which was previously associated with heat-stress in chickens, sub-clinical mastitis in dairy cows and average daily gain in Nellore cattle. Other genes flanking these SNPs had previously been associated with immune and disease-related traits in livestock, as well as growth and feed efficiency. However, the overall genetic architecture for the reaction norm slope appeared to be highly polygenic and the relative effect of the significant SNPs was small, meaning the value of understanding the role of the individual genes for the prediction of breeding values for robustness will likely be negligible.

The chapter also highlighted the potential impact of the different scale-type G×E interactions on the interpretation of variation in the reaction norm slope and its correlation with the intercept. It showed that the correlation between the intercept and slope is interpreted more accurately as change in genetic variance across environments, rather than a correlation between overall performance and robustness. It also showed how this correlation could influence the resulting associations between SNPs and the reaction norm slope and suggested a potential

‘scale-correction’, which used genetic regression to remove variation due to scale-type $G \times E$ interaction in the reaction norm slope.

Chapter four expanded on this idea by demonstrating the ability of the genetic regression, as well as an alternative standardisation method, to remove variation in the breeding values for the reaction norm slope originating from scale-type $G \times E$ interactions. This was achieved using a multi-environment trial dataset from an Australian barley breeding program which had large amounts of both scale-type and rank-type $G \times E$ interactions. We demonstrated the effectiveness of the reaction norm scale-correction methods by leveraging the capacity of factor analytic models to explicitly separate scale and rank-type $G \times E$ interactions. After applying the scale-corrections, the breeding values for the reaction norm slope were highly correlated with the breeding values for robustness from the factor analytic models. The chapter showed that variation in the reaction norm slope needs to be interpreted with reference to the relative contribution of scale-type and rank-type $G \times E$ interactions, because the two types of interaction have very different properties and therefore impact on the potential selection for robustness differently.

In chapter five, factor-analytic models were used to explore $G \times E$ interactions and variation in the robustness of the post-weaning body weight of lambs between 31 discrete environments (flock-years). A factor-analytic model with two common factors was able to approximate the unstructured genetic covariance matrix between all the environments using 85% fewer parameters than what would have been required by an unstructured multi-trait model. We demonstrated that environments could be clustered by their similarity and showed that $G \times E$ interactions were large both between flocks and across years within flocks. Although the breeding values for overall performance across environments were highly correlated between the factor-analytic and the more commonly applied reaction norm models (0.98), the correlation between breeding values for robustness was low (0.16-0.17). The estimates of

heritability, genetic variance, and genetic correlations between environments were also inconsistent between the models, and it was unclear which model was preferred based on the goodness-of-fit tests.

The sixth chapter of this thesis concentrated on the ability of reaction norm breeding values to predict the robustness of a sire's progeny performance. We showed that slope breeding values estimated using reaction norm models were predictive of the phenotypic robustness of progeny across different environments within a research flock. We then verified that the reaction norm breeding values estimated in the research flock were consistent with breeding values estimated independently in the wider industry population recorded by commercial stud breeders. Therefore, selection based on reaction norm breeding values could be used to increase the robustness of body weight gain in Australian sheep to variation in growth environments. Given the findings of this thesis, this general discussion will focus on three main points. These are: 1) the statistical and biological definition of robustness, 2) the choice of model for estimating $G \times E$ interactions and robustness, and 3) the potential of selection for robustness in practice.

7.1.1 The statistical and biological definition of robustness

At the beginning of this thesis, the conventional understanding in both animal and plant genetics was that the reaction norm slope could be interpreted as a breeding value for the robustness of a genotype, and the correlation between the intercept and the slope represented the genetic relationship between overall performance in the average environment and robustness (Knap & Su, 2008; Kusmec et al., 2017; Li et al., 2019; Li & Hermes, 2016; Rauw & Gomez-Raya, 2015). We showed that there was significant genetic variation in the reaction norm slope of Australian sheep for post-weaning body weight to different growth environments using both research and industry datasets. Estimates of the genetic correlation between the reaction norm intercept and slope were positive and ranged between 0.04 and 0.64. There was also significant genetic variation in the reaction norm slope for yield in an Australian

barley population across different yield environments, and it was highly correlated with the variation in the intercept (0.76). Under the conventional framework, we could conclude that in these populations the same genes which increased the overall performance of genotypes also decreased their robustness to environmental variation. Similar conclusions have been made using reaction norms in other species (Knap & Su, 2008; Kolmodin et al., 2002; Mulder et al., 2013). This has been explained at a biological level using resource allocation models, which contend that high-performing individuals partition too much energy into their growth and production at the expense of other important processes, such as the immune system, leaving them vulnerable to environmental variation (Friggens & Van Der Waaij, 2008; Rauw & Gomez-Raya, 2015). Although this is a persuading explanation, the presence of a correlation between the intercept and slope also means that genotypes with negative intercepts of equal magnitude to genotypes with positive intercepts will have equivalently large slopes (Waters, Clark, et al., 2022). That is, their breeding values will be equally as sensitive to environmental variation. Hence, it is important to investigate the properties of the reaction norm slope more deeply before speculating on its biological underpinning.

In reaction norm models, genotypes can vary in their slope because of two types of $G \times E$ interaction: rank-type and scale-type. Scale-type $G \times E$ interactions occur when the genetic variance of a trait changes across different environments (James, 2009). This often occurs because environments with higher trait means also tend to have higher trait variances (Falconer, 1990). An example of a reaction norm impacted only by scale-type $G \times E$ is given in Figure 7.1. It is clear that the genetic correlation between all environments will be unity if scale-type $G \times E$ is the only source of interaction in a population. Therefore, performance in all the environments is genetically the same trait, and the rankings of genotypes (and the optimal selection decisions) will be the same across all environments. However, the differences between genotypes get

larger and or smaller according to the variance. It is important to consider whether this variation in the slope is meaningful when trying to select for increased robustness.

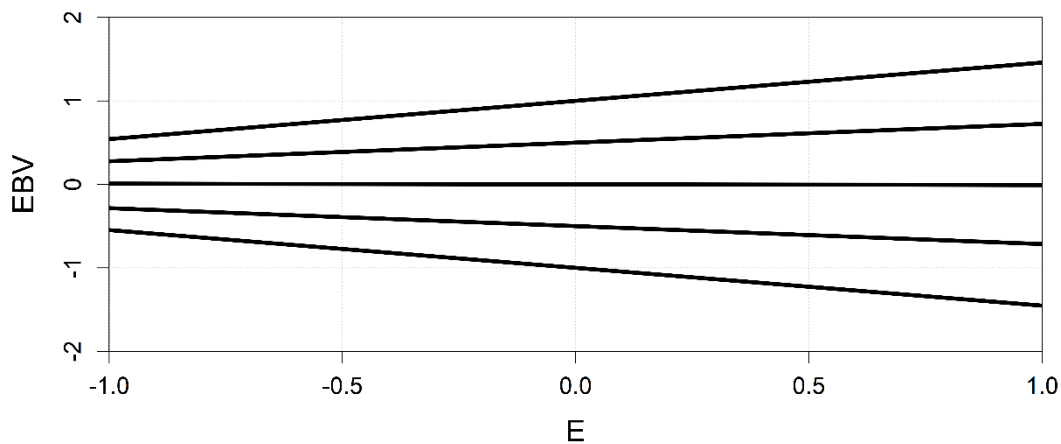


Figure 7. 1 Reaction norm for 5 genotypes (black lines) where the interaction across the environmental covariable, E, is only due to scale-type G×E interaction.

Scale-type G×E does not offer any exploitable genetic variation that is independent of the genetic variation for overall performance because it does not change the relative effect of the underlying genes of traits (and hence the ranking of genotypes) across environments. This is clear in Figure 7.1; the slope is completely dependent on the average breeding value of each genotype across the environments, making it impossible to change the overall performance without changing the slope. Therefore, scale-type G×E is not useful in the selection of robustness in breeding programs. Moreover, the genetic variance, and by extension, scale-type G×E interaction, is a property of the population of genotypes, rather than any individual genotype. Consequently, changes in the breeding value of a genotype when moving into a different environment is, in the case of scale-type G×E interaction, because the variance of the population changes, not because of any attribute intrinsic to the genotype. In contrast, the ranking of a genotype and how it changes across environments (i.e., rank-type G×E) is a property of the individual genotype. This is an important distinction to make in the context of selection decisions, where we are interested in differences between individual genotypes. From

this perspective, it makes sense to distil the relative response of a genotype to different environments by removing the impact of scale-type $G \times E$ on the reaction norm slope. This is further emphasized by the fact that in a reaction norm, a genotype with a high intercept and flat slope will decrease in ranking if the genetic variance increases as the environmental covariable increases (Waters et al., 2022). Therefore, the slope is not an accurate measure of robustness in the presence of scale-type $G \times E$, so it is important to account for it in the analysis of robustness.

In chapter 3 and 4, we established that the genetic correlation between the intercept and slope is a function of scale-type $G \times E$ (Waters et al., 2022; Waters et al., 2023). Consequently, a larger correlation between the intercept and slope implies a greater contribution of scale-type $G \times E$ interaction to the variation in the reaction norm slope, provided that the intercept is placed in the average environment. We used this relationship to propose a genetic regression method to calculate slope breeding values that were independent of the correlation with the intercept (and therefore scale-type $G \times E$). Although we did not know this at the time, the same idea was originally suggested in a two-environment context by Falconer (1990), and later by Telfer et al. (2021). Hence, this thesis represents an extension of the method into a multiple-environment context. Breeding values resulting from the genetic regression capture the change in rank of a genotype across the environmental covariable, which is consistent with how robustness is captured in the factor analytic models used to select genotypes in Australian plant breeding programs (Waters et al., 2023).

Given there are plausible explanations for high overall performance being unfavourably correlated to robustness, it may be unwise to assume that the intercept and slope are uncorrelated using genetic regression. We therefore explored another option to remove scale-type $G \times E$ by directly standardising the breeding values across environments by the estimated genetic variance. The advantage of this approach was that it makes no assumptions based on

the relationship between the intercept and slope and allows the use of higher order polynomials to model the genetic effects. While the results in chapter four indicated that this was less effective at removing scale-type $G \times E$ interactions than the genetic regression method, it could still be useful for the analysis of curvilinear reaction norms (Waters et al., 2023). More work is needed to understand the properties and limitations of the different scale correction methods and how a relationship between overall performance and robustness might manifest in these models.

Previous analyses of the reaction norm slope in both plants and animals have identified pleiotropic genes that were associated with both the intercept and slope (Li et al., 2019; Shi et al., 2021). While this has been interpreted as a common genetic background for overall performance and robustness, this thesis suggests that such findings are result of the correlation between the intercept and slope and hence scale-type $G \times E$ interactions. This was demonstrated in chapter 3, where a QTL on chromosome 11 was significantly associated with both the reaction norm intercept and slope for post-weaning weight. The association with the slope disappeared when the genetic regression was applied. This was because the QTL effects size changed across the environments in proportion to the change in genetic variance. It therefore did not represent a difference in the genetic architecture of post-weaning weight across environments and was only a scale effect. After applying the genetic regression, a new QTL on chromosome 7 was associated with the scale-corrected slope. The relative effect of the QTL changed significantly across the environmental gradient and hence contributed to the re-ranking of genotypes. This suggests the presence of a gene that had an environment-dependent expression. The standardisation method developed later in chapter 4 was also applied to this analysis and revealed very similar results to the genetic regression. It is important for future genomic analyses of the reaction norm slope to consider the use of a scale-correction to better understand how the genetic architecture of traits change across environments.

The biological understanding of robustness is clearer when using a definition that is based on the genetic variation captured by the models used to estimate it. In the context of genotype by environment interactions, robust genotypes have a genetic ranking that is stable across environments. Such genotypes are equipped with genes that enable them to rank consistently across many different environments, and the effect of these genes can be captured using scale-corrected reaction norms and factor-analytic models. This contrasts with the descriptive definitions that have been used to describe reaction norms (Berghof et al., 2019; Doeschl-Wilson et al., 2021; Rauw & Gomez-Raya, 2015), which can be harder to interpret. A potential limitation of the definition proposed in this thesis, however, is that robustness in a single trait does not necessarily imply robustness in other traits, or that a genotype will be robust in a holistic sense (Friggens et al., 2017; Friggens & Van Der Waaij, 2008). This could be further investigated by exploring correlations between other robustness related traits (e.g., worm egg count in sheep, or rust susceptibility in barley). However, it seems likely that the ability of 1) a lamb to gain body weight more reliably across low and high growth environments, and 2) a barley cultivar to yield more reliably in both low and high yielding environments is indicative of a genotype that is robust in a general sense, provided the analysed populations were exposed to a representative sample of environmental factors.

7.1.2 The choice of model for estimating $G \times E$ interactions and robustness

Genotype by environment interactions mean that the genetic variance and/or covariance is heterogeneous between environments. It is generally infeasible to jointly estimate the variances and covariances between all the pairs of environments (e.g., trial $\times$ year, flock $\times$ year or contemporary groups) when there are many discrete environments in the data. Assuming it has some underlying structure could therefore be helpful. The range of genetic models available to capture $G \times E$ interactions differ in how they parameterise the underlying structure. Within these models, there is a direct relationship between the detail at which $G \times E$ interactions are captured,

and the computational and data structure requirements. Large-scale genetic evaluations of livestock that include G×E interactions typically elect a simple interaction model (e.g., sire ×flock) for computational reasons (Golden et al., 2009; Graser et al., 2005; Li et al., 2015). While this allows the genetic correlation between environments to be less than unity, it is uniform between all the pairs of environments and the genetic variance is constant across environments (Smith et al., 2005). This is clearly a limited parameterisation of G×E interactions, and it does not allow the robustness of a genotype to be quantified. The main alternatives to this model were explored in this thesis. These were reaction norm and factor analytic models.

Reaction norm models assume that the G×E interactions acting in a population can be explained by the regression of breeding values across an environmental covariable (EC). While they allow the genetic variance and genetic correlation between environments to vary, the pattern of variance and correlations are solely a function of the chosen EC. This tends to be an oversimplification of the true G×E interaction acting in field populations, where there are many complex and interacting environmental factors that influence the performance of a genotype (van Eeuwijk et al., 2016; Waters et al., 2023). In an attempt to capture these variables into a single EC, we used the popular approach of defining the EC as the mean performance of a group of genotypes within an environment (Carvalho et al., 2019; Kusmec et al., 2017; Lacaze et al., 2009; Via et al., 1995). However, environments with the same EC value are assumed to have the same effect on the genotypes (i.e., a genetic correlation of one, and the same genetic variance), regardless of the underlying environmental factors contributed to the EC value. We found this was too restrictive in the context of modelling G×E interactions in a multi-environment trial in barley (Waters et al., 2023), but it should theoretically be sufficient if there is an identifiable EC that explains most of the G×E interaction. This is often the case in experiments where the EC is known and controlled (e.g., Houshmandfar et al., 2019; Paschke

et al., 2003; Sadras et al., 2012), or the main source of G×E for the trait (Li et al., 2018; Millet et al., 2019).

Several studies have reported that fitting higher-order polynomials can provide a better fit than a linear function (Carvalho et al., 2019; Carvalho Filho et al., 2022). This was also found in chapter 4 of this thesis, where a cubic function was superior to the linear reaction norm based on the model testing criteria (Waters et al., 2023). However, the cubic function appeared to over-fit the data, as the individual curves tended to be highly irregular across the EC. An extreme example of this phenomenon is also demonstrated in de Souza et al. (2020), where a polynomial of order 5 provided the best fit. These examples are probably a symptom of the higher-order regression coefficients attempting to capture trial or farm-specific interactions, rather than the general interactions that were due to the EC. We therefore recommend the use of lower order functions (i.e., linear or quadratic) in field data where the G×E interactions are likely to be complex, even if they provide a poorer fit based on goodness of fit criteria.

Despite their constraints, the assumptions used by reaction norms are computationally and practically convenient. The model requires few parameters to estimate, and the output gives the robustness of genotypes to broadly defined environments (e.g., low vs high production), which is easy to apply in breeding programs. As discussed in chapter 5, the data structure required to fit reaction norms in livestock is more readily available compared to the more complex factor analytic models, since they can exploit environmental variation between contemporary groups within a farm. In addition to the validation study in chapter 6, there is strong evidence that reaction norms can predict the robustness of progeny performance (Nguyen et al., 2016) and increase the accuracy of phenotypic predictions (Mota et al., 2020; Oliveira et al., 2018; Silva et al., 2014) in the presence of G×E interactions. This provides confidence that despite being a generalisation of a complex system, reaction norms capture useful information on G×E interactions and robustness and can offer an advantages over the

models typically applied in genetic evaluations of livestock. If greater flexibility is required to capture the G×E interactions and the appropriate data structure is available, factor analytic models might be a better option.

Factor analytic models approximate the unstructured variance-covariance matrix between environments by assuming it can be explained by the regression of breeding values across ‘latent’ common factors that are estimated from the data. Although this requires significantly more parameters to be estimated compared to the reaction norm models, it is more flexible and less prescriptive (Meyer, 2009), meaning they often perform well in the analysis of multi-environment trial data in plant breeding (Kelly et al., 2007). The models have also been proposed for the analysis of G×E interactions of livestock (Meyer, 2009), but there have only been a few examples of their application (Li et al., 2019; Sae-Lim et al., 2014; Tyrisevä et al., 2011). Methods which derive estimates of robustness based on the regression of genotypes on the latent common factors have also recently been outlined (Smith & Cullis, 2018) and would be interesting to compare with the reaction norm models. As anticipated, the factor analytic models were superior to the reaction norm models based on the AIC and BIC in the multi-environment trial of barley (Waters et al., 2023). Despite this, the reaction norm models provided very similar breeding values for both the overall performance and robustness of genotypes, but only after using the scale-correction methods. When applied to analysis of G×E interactions between flock-years in sheep (chapter 5), the factor analytic models were superior to the reaction norms based on the AIC but poorer based on the BIC. Despite the lack of a clear preference, the factor analytic models provided useful insight into the genetic correlation between different flock-years without forcing a covariance structure based on a prescribed EC. The research demonstrated the potential for factor analytic models to enhance the analysis of G×E interactions in field livestock data, but further investigation of the appropriate data

structure (e.g., genetic linkage between environments and the number of individuals per environment) is required to understand their role.

7.1.3 The potential of selection for robustness in practice

The heritability of a trait is central in predicting the response to selection in a breeding program because it influences the accuracy of genetic predictions (Visscher et al., 2008). Heritability is calculated as the ratio of genetic variance to phenotypic variance and is useful in comparing different traits as it is a dimensionless quantity. The robustness of a genotype cannot be directly measured as a phenotype on a single individual but instead must be estimated through the performance of relatives in different environments. This means that the phenotypic variance and hence the heritability of robustness is not estimable (Sae-Lim et al., 2015), despite some estimates existing in the literature (Knap & Su, 2008). While some alternatives to heritability have been suggested, such as the ‘coheritability’ of the reaction norm slope with the intercept (Sae-Lim et al., 2015), it seems more straightforward to simply quantify the level of exploitable genetic variation in robustness using the units of measurement (e.g., variance in body weight per unit change in growth environment (Waters et al., 2022)). As discussed in the fourth chapter of this thesis, the genetic variation in the reaction norm slope (and hence robustness) after discounting the impact of scale-type G×E is calculated as: $\sigma_{a_1}^{2*} = \sigma_{a_1}^2 \times k$

where $k = \left(1 - \left(\frac{\sigma_{a_0 a_1}}{\sqrt{\sigma_{a_0}^2 \sigma_{a_1}^2}} \right)^2 \right)$, which is the proportion of genetic variation due in the slope

due to rank-type G×E (Falconer, 1990), $\sigma_{a_0}^2$ and $\sigma_{a_1}^2$ are the genetic variances in the intercept and slope, respectively, and $\sigma_{a_0 a_1}$ is the covariance between them. The response to directional selection (R) can then be predicted using $R = i \cdot r \cdot \sigma_{a_1}^{2*}$, where i is the selection intensity and r is the accuracy of the robustness EBVs. However, it is more likely that selection based on robustness EBVs would be stabilising, where a breeding value closer to zero is desirable.

Response to selection is more difficult to calculate in this case, but recently derived selection index theory based on stabilising selection might be informative (Wellmann, 2023). In the context of factor analytic models, the average proportion of additive genetic variance explained by the common factors used to calculate robustness (equivalent to k in Eq. 1) could be multiplied by the overall average additive genetic variance across the environments to obtain a comparable value of σ_{a1}^{2*} . Therefore, it is not necessary to calculate the heritability of robustness for it to be implemented in breeding programs.

Chapter 6 demonstrated that the reaction norm models could successfully predict the robustness of a sires' offspring to gain body weight across different growth environments. This indicates we could apply these models in genetic evaluations of sheep to incorporate robustness into selection decisions. Before implementation of these models, there some additional factors which should be explored, including data structure and the impact of unbalanced information on the estimation of reaction norm breeding values for robustness. The breeding value of a genotype is shrunk depending on the amount of information available to estimate it. A previous study in pigs observed that sires with a wider environmental range had both larger slopes and smaller standard errors (Li & Hermes, 2016), suggesting that estimates of robustness were confounded with the data structure. The impact of data structure on the estimation of breeding values at a given EC level has also been explored (Calus et al., 2004), but they did not explicitly examine the impact on the estimation of slope EBVs for individual genotypes. This question is particularly important given that industry datasets tend to be unbalanced.

More work is also required to understand the economic value of robustness, as well as its correlation with other traits. We demonstrated in chapter 5 of that genetic correlations within the same farm across different years can be much less than unity, meaning that individual farms would likely benefit from the selection of robust genotypes. However, this benefit should be formally quantified through economic modelling to understand its relative importance in a

selection index. If the benefits of good performance in favourable environments outweigh the costs of poor performance in unfavourable environments, it is possible that robustness may not be of economic significance. However, it will also be important to consider other factors such as the welfare implications of poor performance of in unfavourable environments. The level of environmental variability expected within a farm is also an important factor. For example, robustness will be of no value in environments that are consistent or highly predictable, such as in glasshouses (Heuvelink, 2005). In this case, selection could be simply made on the performance of genotypes within the given environment. In addition to the economic value, the correlation of robustness with other traits is also important in a selection index. These can be calculated by directly regressing breeding values for robustness with other EBVs included in the index, as described in Nguyen et al. (2017). With this information, robustness could be included in a selection index which appropriately balances its importance relative to other traits. However, inclusion in a selection index is not necessary for robustness to be utilised in breeding programs.

Outside of selection indexes, breeding values for overall performance and robustness could be directly provided for each relevant trait (e.g., the intercept and slope for post-weaning weight in sheep). This would enable individual farms to understand how the breeding value of a genotype may change depending on the season (unfavourable vs favourable) and make selections according to their attitude towards risk. Additionally, robust genotypes could be created for farms where it is important by tactically mating individuals with high but contrasting sensitivities. If two genotypes with extremely positive and negative slopes (based on a reaction norm) are mated, the resulting progeny will likely be equipped with the range of favourable alleles required for performance across the environmental trajectory. These tailored matings could be designed using programs such as MateSel (Kinghorn et al., 2015).

Information on the robustness of different breeds (or genetic groups) could also be used to inform the selection of animals for a given farm. In chapter 3, a linear regression of the genetic group component across the environmental covariable was tested but was not significantly different from zero. This indicated that while the main effect of genetic groups was significant, the genetic groups did not vary in their effect across environments and hence robustness. The result should be interpreted with caution however, as only the Merino breed was represented as pure-breds while the rest were represented as crosses with Merinos.

Robustness could potentially be increased by utilising heterosis. A fantastic example of this can be found in maize. In the early 20th century, the maize population consisted of many open-pollinated varieties that were highly adapted to local climates (Troyer, 1996). The introduction of hybrid Maize in the 1930's dramatically increased both the yield and adaptability (i.e., robustness) of varieties, such that they could be used over a much greater geographical range. In animal experiments, first-cross hybrids of mice (Ainsworth Harrison, 1968) and beef cattle (Gaughan et al., 1999) have shown superior heat tolerance relative to their pure-bred counterparts. However, there is an overall lack of scientific studies which investigate the relative robustness of cross- and pure-bred animals and hence is suggested as an area for future research.

There are unique challenges to the application of factor analytic models for the routine genetic evaluation of livestock. While plant breeding programs are typically operated in-house by a small number of individuals with training in quantitative genetics, sheep and cattle breeding programs are highly dispersed, with many individual breeders (e.g., >1000, Sheep Genetics, 2022) that collectively contribute to genetic progress (van der Werf & Banks, 2017). In this setting, new technologies and strategies should be clearly communicated for them to be adopted successfully (Gamborg & Sandøe, 2005). The implicit nature of the environmental covariables that are used for the regression of genetic merit in factor analytic models could make their

implementation challenging, as they are hard to interpret and communicate to breeders in straight-forward way. Other problems could arise from how the EBVs for robustness are estimated in the factor analytic models. Since the common factors are data-dependent, the environmental covariates implicit in the common factors would change as new data is added, meaning the breeding values for robustness would represent a different underlying genetic component between analyses. This potential inconsistency would be undesirable for genetic evaluations, making a reaction-norm approach utilising static and interpretable environmental factors more appealing. Furthermore, the genetic evaluations of livestock can involve more than 30 different traits (Boerner & Johnston, 2019; Brown et al., 2018), for which the genetic variances and correlations are assumed to be known. Although factor-analytic models handling several traits across multiple environments have been developed for the analysis of plant breeding data (Bančič et al., 2023), the number of environments and genotypes, and the smaller level of genetic linkage found in livestock data, would likely be prohibitive of such models. It seems most likely the application of factor analytic models to livestock genetics might be best served for the analysis of well-linked populations that have data over a range of environments which cannot easily be capture by one environmental covariable.

7.2 Conclusions

Genotype by environment interaction is a source of phenotypic variation for many traits important to agricultural production. This thesis has shown that variation due to $G \times E$ interactions can be used to identify genotypes that have genetic merit that is more robust to environmental variation. We have highlighted the importance of understanding how the different sources of $G \times E$ interaction (scale-type and rank-type) contribute to the estimation of robustness, and proposed methods to partition their effects in reaction norm models. The thesis demonstrated the application of factor-analytic models for the detailed analysis of $G \times E$ interactions and the robustness of livestock across many discrete environments and contrasted

the merits of these models with reaction norms. Furthermore, we demonstrated the ability of reaction norm models to predict the capacity of a sires' progeny to reliably gain weight across different growth environments in both research and industry data. Several areas for future research were discussed, including the economic value of robustness, the optimal design of data structures for its estimation, the implementation of multiple environment testing and methods for its inclusion in genetic evaluations.

Appendix A

Table A3. 1 Breed composition of the INF and RF data used to estimate the contemporary group effects (a) and the reaction norm models (b)

Breed	a	b
Merino	60.8	60.3
Border Leicester	9.1	9.7
Poll Dorset	10.6	11.8
Suffolk	3.3	3.2
White Suffolk	3.6	4.1
Research	4.1	2.5
Other	8.5	8.4

Table A3. 2 Descriptive statistics for post-weaning growth rate (PWGR) used to estimate contemporary group effects and the post-weaning weight (PWWT) before and after removing non-genotyped animals that were eliminated the reaction norm analysis.

Trait	Male	Female	Mean	SD	Min	Max
PWGR (g/day)	16 898	16 875	117.7	60.3	-129.0	365.71
PWWT (kg) ^a	16 898	16 875	43.4	9.4	9.0	79.5
PWWT (kg) ^b	10 585	11 928	44.2	8.7	13.2	78.8

^a Summary using all animals in the data. ^b Summary using genotyped animals used for the reaction norm analysis.

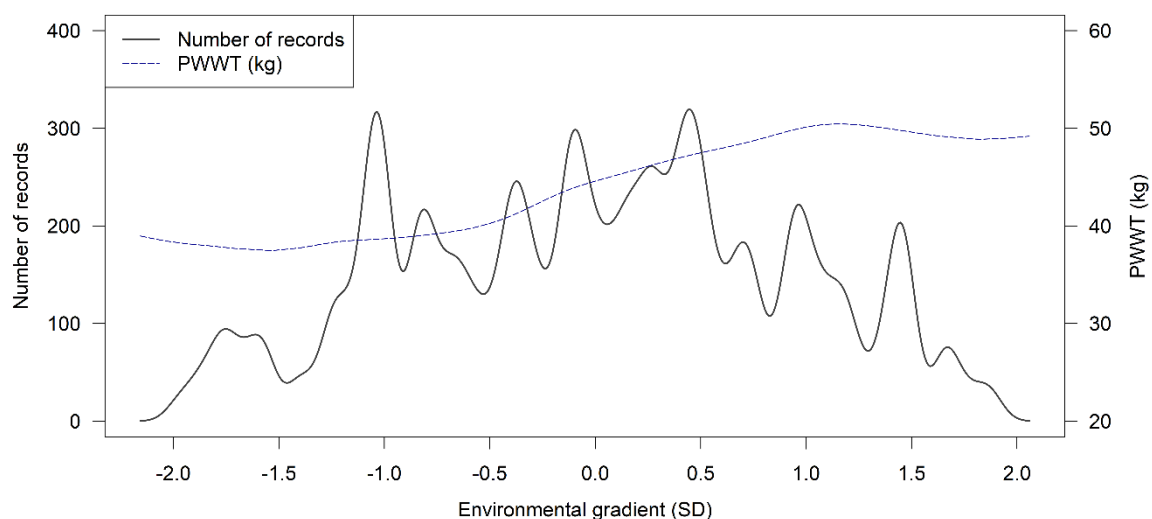


Figure A3. 1 Distribution of animals and mean post-weaning weight across the environmental gradient.

Table A3. 3 Approximate number of reference animals used for imputation for each SNP panel

SNP panel	Number of reference animals
50k 1	183 000
50k 2	88 000
15k 1	62 000
15k 2	15 000
12k	14 000

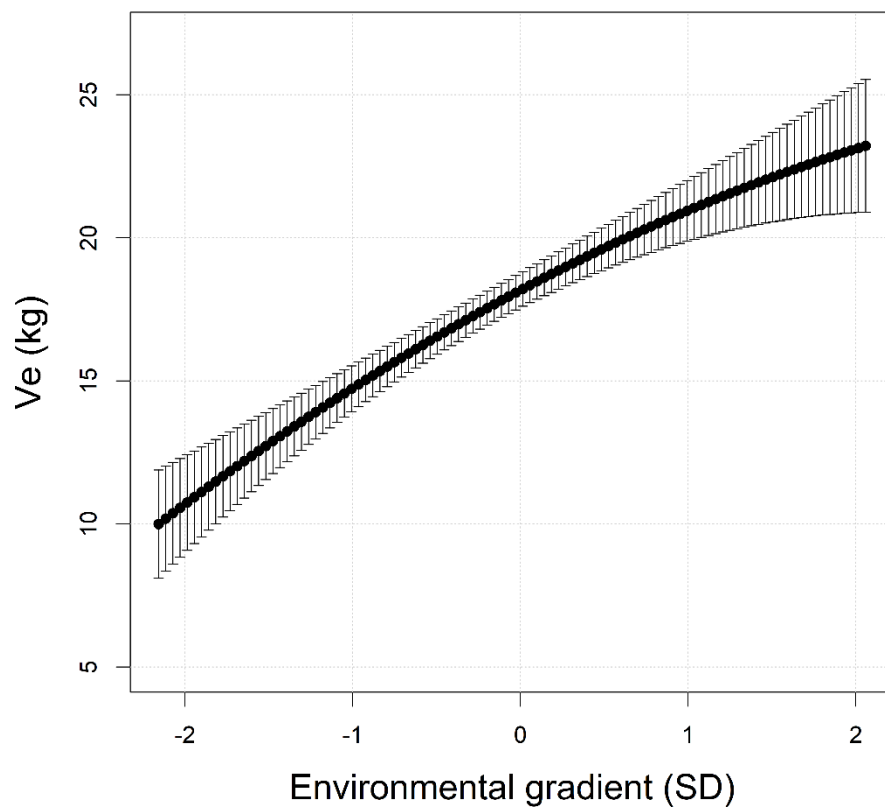


Figure A3. 2 Residual variance along the environmental gradient for the RNM-HET.

Table A3. 4 Additive genetic variance for each environment (low, average and high growth environments) in the three different models

Environment	Model ^a		
	RNM_HOM	RNM_HET	MT
Low	5.46 (0.45)	8.12 (0.60)	9.31 (0.97)
Average	7.45 (0.32)	7.00 (0.30)	7.04 (0.49)
High	15.14 (0.76)	11.25 (0.80)	13.57 (1.35)

Table A3. 5 Heritability for each growth environment (low, average and high) in the three different models

Environment	Model ^a		
	RNM_HOM	RNM_HET	MT
Low	0.24 (0.02)	0.39 (0.03)	0.31 (0.02)
Average	0.30 (0.01)	0.28 (0.01)	0.23 (0.01)
High	0.46 (0.01)	0.34 (0.02)	0.29 (0.02)

Table A3. 6 Genetic correlations between each growth environment (low, average and high) in the three different models

Environment	Model ^a		
	RNM_HOM	RNM_HET	MT
Low vs average	0.77	0.81	0.62
Average vs high	0.92	0.88	0.81
Low vs high	0.49	0.44	0.67

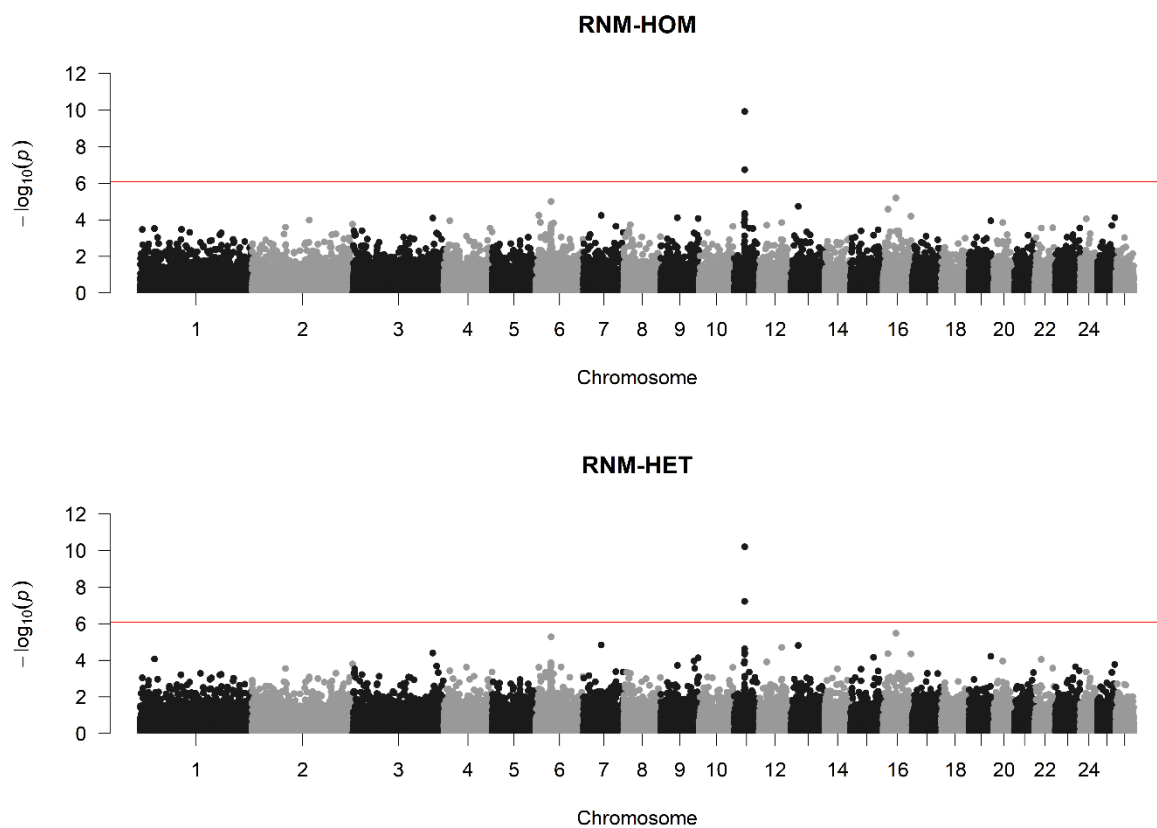


Figure A3. 3 Manhattan plots for the intercept in the RNM-HOM and RNM-HET.

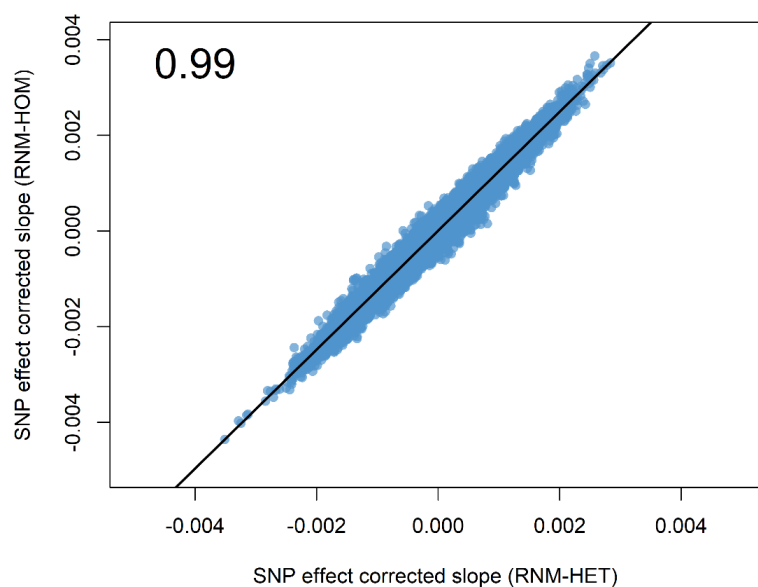


Figure A3. 4 Correlation between SNP effects for scale-corrected slope from the RNM-HOM and RNM-HE

Appendix B

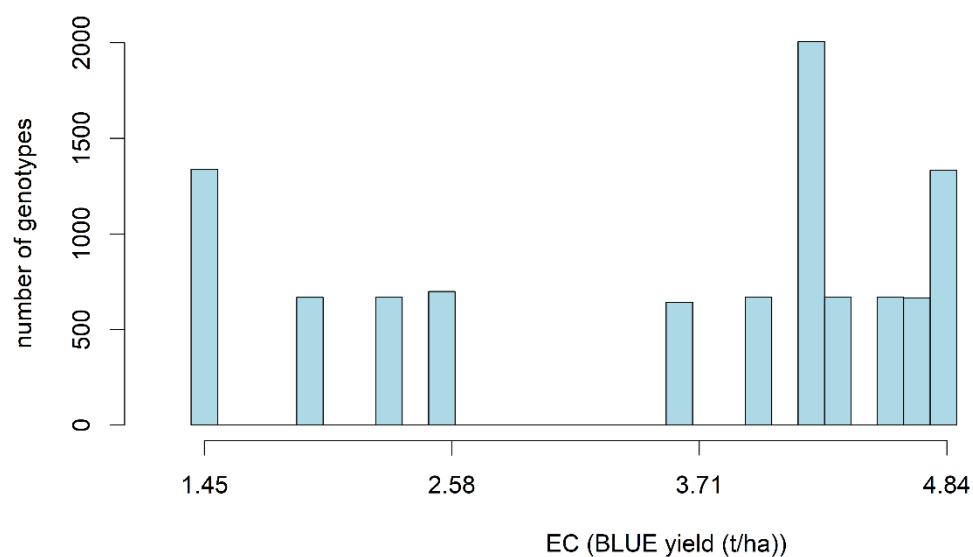


Figure A4. 1 The distribution of genotypes along the EC

Table A4. 1 The percentage of genetic variance explained by the first (v_{a1}) and second (v_{a2}) common factor for each environment in FA-2.

Environment	v_{a1}	v_{a2}	Total ($v_{a1} + v_{a2}$)
1	80.61	6.92	87.53
2	88.87	3.69	92.56
3	8.81	24.23	33.04
4	0.41	93.59	94.00
5	76.00	0.68	76.68
6	88.62	1.31	89.93
7	0.86	99.14	100
8	23.85	18.3	42.15
9	68.43	18.24	86.67
10	32.68	19.21	51.89
11	50.16	3.90	54.06
12	62.73	12.55	75.28
13	26.02	2.66	28.68
14	81.21	4.82	86.03
15	69.22	24.31	93.53
Average	50.56	22.24	72.80

Table A4. 2 Variance (diagonal, bold), covariance (lower triangle) and correlations (upper triangle, italicised) for the random regression coefficients in RN-L and RN-NL.

Coefficients	RN-L			
	a₀	a₁	a₂	a₃
a₀	0.0458	<i>0.76</i>	-	-
a₁	0.0199	0.0151	-	-
a₂	-	-	-	-
a₃	-	-	-	-
Coefficients	RN-NL			
	a₀	a₁	a₂	a₃
a₀	0.0675	<i>0.71</i>	<i>-0.75</i>	<i>-0.68</i>
a₁	0.0291	0.0260	<i>-0.31</i>	<i>-0.87</i>
a₂	-0.0212	-0.0054	0.0119	<i>0.46</i>
a₃	-0.0091	-0.0072	0.0026	0.0026

Table A4. 3 Loadings of the common factors to each environment for FA-2 and RR-2. Negative loadings are highlighted in bold.

Environment	FA-2		RR-2	
	$\tilde{\lambda}_{a_1}^*$	$\tilde{\lambda}_{a_2}^*$	$\tilde{\lambda}_{a_1}^*$	$\tilde{\lambda}_{a_2}^*$
1	0.415	-0.122	0.336	-0.123
2	0.325	0.066	0.304	0.049
3	0.036	0.06	0.051	0.078
4	-0.013	0.192	0.001	0.176
5	0.48	-0.045	0.412	-0.117
6	0.513	-0.062	0.499	-0.071
7	-0.032	0.343	-0.022	0.27
8	0.098	0.086	0.048	0.055
9	0.316	0.163	0.333	0.107
10	0.073	0.056	0.08	0.023
11	0.296	0.083	0.329	0.113
12	0.291	-0.13	0.206	-0.099
13	0.175	-0.056	0.154	0.143
14	0.29	0.071	0.256	0.052
15	0.235	0.139	0.224	0.091

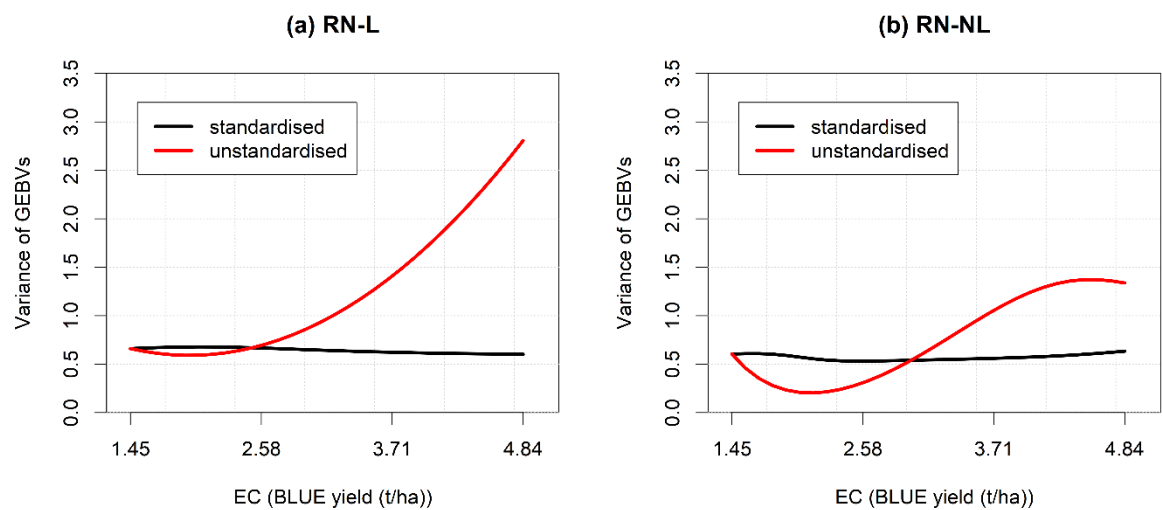


Figure A4. 2 Variance of the GEBVs across the EC before and after standardisation in RN-L and RN-NL.

Chapter 8. Consolidated Reference List

- Agarwala, R., Barrett, T., Beck, J., Benson, D. A., Bollin, C., Bolton, E., Bourexis, D., Brister, J. R., Bryant, S. H., Canese, K., Cavanaugh, M., Charowhas, C., Clark, K., Dondoshansky, I., Feolo, M., Fitzpatrick, L., Funk, K., Geer, L. Y., Gorelenkov, V., ... Zbicz, K. (2018). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 46(D1), D8–D13. <https://doi.org/10.1093/nar/gkx1095>
- Aguilar, I., Misztal, I., Johnson, D. L., Legarra, A., Tsuruta, S., & Lawlor, T. J. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, 93(2). <https://doi.org/10.3168/jds.2009-2730>
- Aoun, M., Carter, A., Thompson, Y. A., Ward, B., & Morris, C. F. (2021). Environment characterization and genomic prediction for end-use quality traits in soft white winter wheat. *Plant Genome*, 14(3). <https://doi.org/10.1002/tpg2.20128>
- Bahlo, C., Dahlhaus, P., Thompson, H., & Trotter, M. (2019). The role of interoperable data standards in precision livestock farming in extensive livestock systems: A review. In *Computers and Electronics in Agriculture* 156 <https://doi.org/10.1016/j.compag.2018.12.007>
- Bančič, J., Ovenden, B., Gorjanc, G., & Tolhurst, D. J. (2023). Genomic selection for genotype performance and stability using information on multiple traits and multiple environments. *Theoretical and Applied Genetics*, 136(5), 104. <https://doi.org/10.1007/s00122-023-04305-1>
- Bapela, T., Shimelis, H., Tsilo, T. J., & Mathew, I. (2022). Genetic Improvement of Wheat for Drought Tolerance: Progress, Challenges and Opportunities. In *Plants* 11(10). <https://doi.org/10.3390/plants11101331>
- Basford, K. E., & Cooper, M. (1998). Genotype x environment interactions and some considerations of their implications for wheat breeding in Australia. In *Australian Journal of Agricultural Research* (Vol. 49, Issue 2). <https://doi.org/10.1071/A97035>
- Berghof, T. V. L., Poppe, M., & Mulder, H. A. (2019). Opportunities to improve resilience in animal breeding programs. *Frontiers in Genetics*, 10 <https://doi.org/10.3389/fgene.2018.00692>

- Boerner, V., & Johnston, D. J. (2019). More animals than markers: A study into the application of the single step T-BLUP model in large-scale multi-trait Australian Angus beef cattle genetic evaluation. *Genetics Selection Evolution*, 51(1). <https://doi.org/10.1186/s12711-019-0499-x>
- Bradbury, L. M. T., Fitzgerald, T. L., Henry, R. J., Jin, Q., & Waters, D. L. E. (2005). The gene for fragrance in rice. *Plant Biotechnology Journal*, 3(3), 363–370. <https://doi.org/10.1111/j.1467-7652.2005.00131.x>
- Bressan, M. C., Rossato, L. V., Rodrigues, E. C., Alves, S. P., Bessa, R. J. B., Ramos, E. M., & Gama, L. T. (2011). Genotype $\times$ environment interactions for fatty acid profiles in *Bos indicus* and *Bos taurus* finished on pasture or grain. *Journal of Animal Science*, 89(1). <https://doi.org/10.2527/jas.2009-2672>
- Brown, D. J., Huisman, A. E., Swan, A. A., Graser, H.-U., Woolaston, R. R., Ball, A. J., Atkins, K. D., & Banks, R. G. (2007). Genetic Evaluation for the Australian Sheep Industry. *Proc. Assoc. Advmt. Anim. Breed. Genet*, 17.
- Brown, D. J., Swan, A. A., Boerner, V., Li, L., Gurman, P. M., McMillan, A. J., van der Werf, J. H. J., Chandler, H. R., Tier, B., & Banks, R. G. (2018). Single-step genetic evaluations in the Australian sheep industry. *Proceedings of the World Congress on Genetics Applied to Livestock Production*, 11(February).
- Brown, M. A., Brown, A. H., Jackson, W. G., & Miesner, J. R. (1993). Genotype $\times$ environment interactions in postweaning performance to yearling in Angus, Brahman, and reciprocal-cross calves. *Journal of Animal Science*, 71(12). <https://doi.org/10.2527/1993.71123273x>
- Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A One-Penny Imputed Genome from Next-Generation Reference Panels. *American Journal of Human Genetics*, 103(3), 338–348. <https://doi.org/10.1016/j.ajhg.2018.07.015>
- Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J., & Thompson, R. (2018). ASReml-R Reference Manual Version 4. *ASReml-R Reference Manual*.
- Calderini, D. F., & Slafer, G. A. (1999). Has yield stability changed with genetic improvement of wheat yield? *Euphytica*, 107(1), 51–59. <https://doi.org/10.1023/A:1003579715714>

- Calus, M. P. L., Bijma, P., & Veerkamp, R. F. (2004). Effects of data structure on the estimation of covariance functions to describe genotype by environment interactions in a reaction norm model. *Genetics Selection Evolution*, 36(5). <https://doi.org/10.1051/gse:2004013>
- Calus, M. P. L., Groen, A. F., & De Jong, G. (2002). Genotype x environment interaction for protein yield in Dutch dairy cattle as quantified by different models. *Journal of Dairy Science*, 85(11). [https://doi.org/10.3168/jds.S0022-0302\(02\)74399-3](https://doi.org/10.3168/jds.S0022-0302(02)74399-3)
- Calus, M. P. L., Janss, L. L. G., & Veerkamp, R. F. (2006). Genotype by environment interaction for somatic cell score across bulk milk somatic cell count and days in milk. *Journal of Dairy Science*, 89(12). [https://doi.org/10.3168/jds.S0022-0302\(06\)72533-4](https://doi.org/10.3168/jds.S0022-0302(06)72533-4)
- Calus, M. P. L., & Veerkamp, R. F. (2003). Estimation of environmental sensitivity of genetic merit for milk production traits using a random regression model. *Journal of Dairy Science*, 86(11). [https://doi.org/10.3168/jds.S0022-0302\(03\)73982-4](https://doi.org/10.3168/jds.S0022-0302(03)73982-4)
- Carabaño, M. J., Wade, K. M., & Van Vleck, L. D. (1990). Genotype by Environment Interactions for Milk and Fat Production Across Regions of the United States. *Journal of Dairy Science*, 73(1). [https://doi.org/10.3168/jds.S0022-0302\(90\)78661-4](https://doi.org/10.3168/jds.S0022-0302(90)78661-4)
- Cardoso, F. F., & Tempelman, R. J. (2012). Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. *Journal of Animal Science*, 90(7). <https://doi.org/10.2527/jas.2011-4333>
- Carvalho, R., Costilla, R., Neves, H. H. R., Albuquerque, L. G., Moore, S., & Hayes, B. J. (2019). Unraveling genetic sensitivity of beef cattle to environmental variation under tropical conditions. *Genetics Selection Evolution*, 51(1). <https://doi.org/10.1186/s12711-019-0470-x>
- Carvalho Filho, I., Silva, D. A., Teixeira, C. S., Silva, T. L., Mota, L. F. M., Albuquerque, L. G., & Carvalho, R. (2022). Heteroscedastic Reaction Norm Models Improve the Assessment of Genotype by Environment Interaction for Growth, Reproductive, and Visual Score Traits in Nellore Cattle. *Animals*, 12(19). <https://doi.org/10.3390/ani12192613>
- Chaudhary, D., Swati, Nagar, K., & Dhyani, R. (2022). Genetic analysis of yield and its attributes in bread wheat (*Triticum aestivum* L. em. Thell) under irrigated and rainfed conditions. *Euphytica*, 218(9). <https://doi.org/10.1007/s10681-022-03080-2>

- Cheng, Z., Buggiotti, L., Salavati, M., Marchitelli, C., Palma-Vera, S., Wylie, A., Takeda, H., Tang, L., Crowe, M. A., Wathes, D. C., Crowe, M., McLoughlin, N., Fahey, A., Matthews, E., Santoro, A., Byrne, C., Rudd, P., O'Flaherty, R., Hallinan, S., ... Palma-Vera, S. (2021). Global transcriptomic profiles of circulating leucocytes in early lactation cows with clinical or subclinical mastitis. *Molecular Biology Reports*, 4611–4623. <https://doi.org/10.1007/s11033-021-06494-8>
- Cheruiyot, E. K., Nguyen, T. T. T., Haile-Mariam, M., Cocks, B. G., Abdelsayed, M., & Pryce, J. E. (2020). Genotype-by-environment (temperature-humidity) interaction of milk production traits in Australian Holstein cattle. *Journal of Dairy Science*, 103(3). <https://doi.org/10.3168/jds.2019-17609>
- Clark, S. A., Hickey, J. M., & van der Werf, J. H. J. (2011). Different models of genetic variation and their effect on genomic evaluation. *Genetics Selection Evolution*, 43(1), 18. <https://doi.org/10.1186/1297-9686-43-18>
- Cooper, M., Voss-Fels, K. P., Messina, C. D., Tang, T., & Hammer, G. L. (2021). Tackling $G \times E \times M$ interactions to close on-farm yield-gaps: creating novel pathways for crop improvement by predicting contributions of genetics and management to crop productivity. In *Theoretical and Applied Genetics* 134(6). <https://doi.org/10.1007/s00122-021-03812-3>
- Crossa, J., Montesinos-López, O. A., Pérez-Rodríguez, P., Costa-Neto, G., Fritsche-Neto, R., Ortiz, R., Martini, J. W. R., Lillemo, M., Montesinos-López, A., Jarquin, D., Breseghello, F., Cuevas, J., & Rincent, R. (2022). Genome and Environment Based Prediction Models and Methods of Complex Traits Incorporating Genotype $\times$ Environment Interaction. In *Methods in Molecular Biology* 2467. https://doi.org/10.1007/978-1-0716-2205-6_9
- Cuevas, J., Crossa, J., Montesinos-López, O. A., Burgueño, J., Pérez-Rodríguez, P., & de los Campos, G. (2017). Bayesian genomic prediction with genotype $\times$ environment interaction kernel models. *G3: Genes, Genomes, Genetics*, 7(1). <https://doi.org/10.1534/g3.116.035584>
- Cullis, B. R., Jefferson, P., Thompson, R., & Smith, A. B. (2014). Factor analytic and reduced animal models for the investigation of additive genotype-by-environment interaction in outcrossing plant species with application to a *Pinus radiata* breeding programme. *Theoretical and Applied Genetics*, 127(10). <https://doi.org/10.1007/s00122-014-2373-0>

- Daltro, D. dos S., Ambrosini, D. P., Negri, R., da Silva, M. V. G. B., & Cobuci, J. A. (2022). Reaction norm model to describe environmental sensitivity in Girolando cattle. *Livestock Science*, 263. <https://doi.org/10.1016/j.livsci.2022.105025>
- de Souza, M. H., Pereira Júnior, J. D., Steckling, S. D. M., Mencalha, J., Dias, F. dos S., Rocha, J. R. do A. S. de C., Carneiro, P. C. S., & Carneiro, J. E. de S. (2020). Adaptability and stability analyses of plants using random regression models. *PLOS ONE*, 15(12), e0233200-. <https://doi.org/10.1371/journal.pone.0233200>
- Derner, J. D., Hunt, L., Filho, K. E., Ritten, J., Capper, J., & Han, G. (2017). *Livestock Production Systems*. https://doi.org/10.1007/978-3-319-46709-2_10
- Dias, K. O. D. G., Gezan, S. A., Guimarães, C. T., Parentoni, S. N., Guimarães, P. E. de O., Carneiro, N. P., Portugal, A. F., Bastos, E. A., Cardoso, M. J., Anoni, C. de O., de Magalhães, J. V., de Souza, J. C., Guimarães, L. J. M., & Pastina, M. M. (2018). Estimating genotype $\times$ environment interaction for and genetic correlations among drought tolerance traits in maize via factor analytic multiplicative mixed models. *Crop Science*, 58(1). <https://doi.org/10.2135/cropsci2016.07.0566>
- Ding, R., Yang, M., Wang, X., Quan, J., Zhuang, Z., Zhou, S., Li, S., Xu, Z., Zheng, E., Cai, G., Liu, D., Huang, W., Yang, J., & Wu, Z. (2018). Genetic architecture of feeding behavior and feed efficiency in a Duroc pig population. *Frontiers in Genetics*, 9(JUN), 1–11. <https://doi.org/10.3389/fgene.2018.00220>
- Doeschl-Wilson, A., Knap, P. W., Opriessnig, T., & More, S. J. (2021). Review: Livestock disease resilience: from individual to herd level. *Animal* 15(1) <https://doi.org/10.1016/j.animal.2021.100286>
- Dominik, S., Newton, J. E., Hayes, B. J., & van der Werf, J. H. J. (2013). Gene by environment interaction for reproduction traits in Australian sheep. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*, 364, 1–186. <https://doi.org/10.1002/lary.21679>
- Donnelly, J. R., McKinney, G. T., & Morley, F. H. W. (1985). The productivity of breeding ewes grazing on lucerne or grass and clover pastures on the tablelands of southern Australia. 4. Lamb growth. *Australian Journal of Agricultural Research*, 36(3). <https://doi.org/10.1071/AR9850469>

- Ducrocq, V., Cadet, A., Patry, C., van der Westhuizen, L., van Wyk, J. B., & Naser, F. W. C. (2022). Two approaches to account for genotype-by-environment interactions for production traits and age at first calving in South African Holstein cattle. *Genetics Selection Evolution*, 54(1). <https://doi.org/10.1186/s12711-022-00735-5>
- Duijvesteijn, N., Bolormaa, S., Daetwyler, H. D., & Werf, J. H. J. Van Der. (2018). Genomic prediction of the polled and horned phenotypes in Merino sheep. *Genetics Selection Evolution*, 50(28), 1–11. <https://doi.org/10.1186/s12711-018-0398-6>
- Eberhart, S. A., & Russell, W. A. (1966). Stability Parameters for Comparing Varieties 1 . *Crop Science*, 6(1). <https://doi.org/10.2135/cropsci1966.0011183x000600010011x>
- Eichi, V. R., Okamoto, M., Garnett, T., Eckermann, P., Darrier, B., Riboni, M., & Langridge, P. (2020). Strengths and weaknesses of national variety trial data for multi-environment analysis: A case study on grain yield and protein content. *Agronomy*, 10(5). <https://doi.org/10.3390/agronomy10050753>
- Falconer, D. S. (1952). The Problem of Environment and Selection. *The American Naturalist*, 86(830). <https://doi.org/10.1086/281736>
- Falconer, D. S. (1990). Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. *Genetical Research*, 56(1), 57–70. <https://doi.org/10.1017/S0016672300028883>
- Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics* (4th ed). Pearson Education Limited.
- Ferrante, A., Cullis, B. R., Smith, A. B., & Able, J. A. (2021). A Multi-Environment Trial Analysis of Frost Susceptibility in Wheat and Barley Under Australian Frost-Prone Field Conditions. *Frontiers in Plant Science*, 12. <https://doi.org/10.3389/fpls.2021.722637>
- Fikse, W. F., Rekaya, R., & Weigel, K. A. (2003). Genotype x environment interaction for milk production in guernsey cattle. *Journal of Dairy Science*, 86(5). [https://doi.org/10.3168/jds.S0022-0302\(03\)73768-0](https://doi.org/10.3168/jds.S0022-0302(03)73768-0)
- Finlay, K. W., & Wilkinson, G. N. (1963). The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research*, 14(6), 742–754. <https://doi.org/10.1071/AR9630742>

- Fischer, T. M., Gilmour, A. R., & Werf, J. (2004). Computing approximate standard errors for genetic parameters derived from random regression models fitted by average information REML. *Genetics Selection Evolution*, 36(3), 363–369. <https://doi.org/10.1186/1297-9686-36-3-363>
- Friggens, N. C., Blanc, F., Berry, D. P., & Puillet, L. (2017). Review: Deciphering animal robustness. A synthesis to facilitate its use in livestock breeding and management. In *Animal* 11(12) <https://doi.org/10.1017/S175173111700088X>
- Friggens, N. C., & Van Der Waaij, E. H. (2008). Modelling of resource allocation patterns. In *Resource Allocation Theory Applied to Farm Animal Production*. <https://doi.org/10.1079/9781845933944.0302>
- Fritsche-Neto, R., Miranda, G. V., DeLima, R. O., & Souza, H. N. de. (2010). Factor analysis and SREG GGE biplot for the genotype $\times$ environment interaction stratification in maize. *Ciência Rural*, 40(5). <https://doi.org/10.1590/s0103-84782010000500007>
- Gamborg, C., & Sandøe, P. (2005). Sustainability in farm animal breeding: a review. *Livestock Production Science*, 92(3), 221–231. <https://doi.org/10.1016/J.LIVPRODSCI.2004.08.010>
- Gaughan, J. B., Mader, T. L., Holt, S. M., Josey, M. J., & Rowan, K. J. (1999). Heat tolerance of Boran and Tuli crossbred steers. *Journal of Animal Science*, 77(9), 2398–2405. <https://doi.org/10.2527/1999.7792398x>
- Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J., & Thompson, R. (2021). ASReml User Guide Release 4.2 Functional Specification. In *VSN International Ltd. Hemel Hempstead* (Issue C). www.vsn.co.uk.
- Golden, B. L., Garrick, D. J., & Benyshek, L. L. (2009). Milestones in beef cattle genetic evaluation. *Journal of Animal Science*, 87(14 Suppl). <https://doi.org/10.2527/jas.2008-1430>
- Gondro, C. (2015). *Primer to Analysis of Genomic Data Using R*. Springer. Cham.
- Gourdine, J. L., Riquet, J., Rosé, R., Pouillet, N., Giorgi, M., Billon, Y., Renaudeau, D., & Gilbert, H. (2019). Genotype by environment interactions for performance and thermoregulation responses in growing pigs. *Journal of Animal Science*, 97(9). <https://doi.org/10.1093/jas/skz245>

- Graser, H.-U., Tier, B., Johnston, D. J., & Barwick, S. A. (2005). Genetic evaluation for the beef industry in Australia. *Australian Journal of Experimental Agriculture*, 45(8), 913–921. <https://doi.org/10.1071/EA05075>
- Guy, S. Z. Y., Li, L., Thomson, P. C., & Hermes, S. (2019). Reaction norm analysis of pig growth using environmental descriptors based on alternative traits. *Journal of Animal Breeding and Genetics*, 136(3), 153–167. <https://doi.org/10.1111/jbg.12388>
- Haile-Mariam, M., Carrick, M. J., & Goddard, M. E. (2008). Genotype by environment interaction for fertility, survival, and milk production traits in Australian dairy cattle. *Journal of Dairy Science*, 91(12). <https://doi.org/10.3168/jds.2008-1084>
- Hardner, C., Dieters, M., DeLacy, I., Neal, J., Fletcher, S., Dale, G., & Basford, K. (2011). Identifying deployment zones for *Eucalyptus camaldulensis* x *E. globulus* and x *E. grandis* hybrids using factor analytic modelling of genotype by environment interaction. *Australian Forestry*, 74(1), 30–35. <https://doi.org/10.1080/00049158.2011.10676343>
- Harrison, G.A. (1962). Heterosis and adaptability in the heat tolerance of mice. *Genetics*, 47(4), 427–434, <https://doi.org/10.1093/genetics/47.4.427>
- Hay, E. H., & Roberts, A. (2018). Genome-wide association study for carcass traits in a composite beef cattle breed. *Livestock Science*, 213(April), 35–43. <https://doi.org/10.1016/j.livsci.2018.04.018>
- Hayes, B. J., Bowman, P. J., Chamberlain, A. J., Savin, K., van Tassell, C. P., Sonstegard, T. S., & Goddard, M. E. (2009). A Validated Genome Wide Association Study to Breed Cattle Adapted to an Environment Altered by Climate Change. *PLOS ONE*, 4(8), e6676. <https://doi.org/10.1371/journal.pone.0006676>
- Hayes, B. J., Carrick, M., Bowman, P., & Goddard, M. E. (2003). Genotype × Environment Interaction for Milk Production of Daughters of Australian Dairy Sires from Test-Day Records. *Journal of Dairy Science*, 86(11), 3736–3744. [https://doi.org/10.3168/JDS.S0022-0302\(03\)73980-0](https://doi.org/10.3168/JDS.S0022-0302(03)73980-0)
- Hayes, B. J., Daetwyler, H. D., & Goddard, M. E. (2016). Models for genome × Environment interaction: Examples in livestock. *Crop Science*, 56(5). <https://doi.org/10.2135/cropsci2015.07.0451>

- Hazard, D., Bouix, J., Chassier, M., Delval, E., Foulquié, D., Fassier, T., Bourdillon, Y., François, D., & Boissy, A. (2016). Genotype by environment interactions for behavioral reactivity in sheep. *Journal of Animal Science*, 94(4). <https://doi.org/10.2527/jas.2015-0277>
- Hébert, Y., Guingo, E., & Loudet, O. (2001). The response of root/shoot partitioning and root morphology to light reduction in maize genotypes. *Crop Science*, 41(2). <https://doi.org/10.2135/cropsci2001.412363x>
- Hedhly, A., Hormaza, J. I., & Herrero, M. (2005). Influence of genotype-temperature interaction on pollen performance. *Journal of Evolutionary Biology*, 18(6). <https://doi.org/10.1111/j.1420-9101.2005.00939.x>
- Henderson, C. R. (1976). A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values. *Biometrics*, 32(1). <https://doi.org/10.2307/2529339>
- Heuvelink, E. (2005). Tomatoes, Crop production science in horticulture series. In *Crop production science in horticulture series*, 1.
- Hollema, B. L., Bijma, P., & van der Werf, J. H. J. (2018). Sensitivity of the breeding values for growth rate and worm egg count to environmental worm burden in Australian Merino sheep. *Journal of Animal Breeding and Genetics*, 135(5). <https://doi.org/10.1111/jbg.12349>
- Horan, B., Dillon, P., Faverdin, P., Delaby, L., Buckley, F., & Rath, M. (2005). The interaction of strain of Holstein-Friesian cows and pasture-based feed systems on milk yield, body weight, and body condition score. *Journal of Dairy Science*, 88(3). [https://doi.org/10.3168/jds.S0022-0302\(05\)72790-9](https://doi.org/10.3168/jds.S0022-0302(05)72790-9)
- Houshmandfar, A., Rebetzke, G. J., Lawes, R., & Tausz, M. (2019). Grain yield responsiveness to water supply in near-isogenic reduced-tillering wheat lines – An engineered crop trait near its upper limit. *European Journal of Agronomy*, 102. <https://doi.org/10.1016/j.eja.2018.11.003>
- Ismael, A., Strandberg, E., Berglund, B., Kargo, M., Fogh, A., & Løvendahl, P. (2016). Genotype by environment interaction for the interval from calving to first insemination

- with regard to calving month and geographic location in Holstein cows in Denmark and Sweden. *Journal of Dairy Science*, 99(7). <https://doi.org/10.3168/jds.2015-10820>
- Iung, L. H. de S., Carvalheiro, R., Neves, H. H. de R., & Mulder, H. A. (2020). Genetics and genomics of uniformity and resilience in livestock and aquaculture species: A review. In *Journal of Animal Breeding and Genetics* 137(3). <https://doi.org/10.1111/jbg.12454>
- James, J. W. (1961). Selection in two environments. *Heredity*, 16(2). <https://doi.org/10.1038/hdy.1961.17>
- James, J. W. (2009). Genotype by Environment Interaction in Farm Animals. In *Adaptation and Fitness in Animal Populations*. https://doi.org/10.1007/978-1-4020-9005-9_10
- Jang, S., Lourenco, D., & Miller, S. (2022). Inclusion of sire by herd interaction effect in the genomic evaluation for weaning weight of American Angus. *Journal of Animal Science*, 100(3). <https://doi.org/10.1093/jas/skac057>
- Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., Piraux, F., Guerreiro, L., Pérez, P., Calus, M., Burgueño, J., & de los Campos, G. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and Applied Genetics*, 127(3). <https://doi.org/10.1007/s00122-013-2243-1>
- Jeyaruban, M. G., Johnston, D. J., & Graser, H. U. (2009). Estimation of genotype x environment interactions for growth, fatness and reproductive traits in Australian Angus cattle. *Animal Production Science*, 49(1). <https://doi.org/10.1071/EA08098>
- Keeble-Gagnère, G., Pasam, R., Forrest, K. L., Wong, D., Robinson, H., Godoy, J., Rattey, A., Moody, D., Mullan, D., Walmsley, T., Daetwyler, H. D., Tibbits, J., & Hayden, M. J. (2021). Novel Design of Imputation-Enabled SNP Arrays for Breeding and Research Applications Supporting Multi-Species Hybridization. *Frontiers in Plant Science*, 12. <https://doi.org/10.3389/fpls.2021.756877>
- Kelly, A. M., Smith, A. B., Eccleston, J. A., & Cullis, B. R. (2007). The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Science*, 47(3). <https://doi.org/10.2135/cropsci2006.08.0540>

- Kempton, R. A. (1984). The use of biplots in interpreting variety by environment interactions. *The Journal of Agricultural Science*, 103(1), 123–135. <https://doi.org/10.1017/S0021859600043392>
- Kerr, R. J., Dutkowski, G. W., Apiolaza, L. A., McRae, T. A., & Tier, B. (2002). Developing a genetic evaluation system for forest tree improvement – the making of TREEPLAN®. *7th World Congress on Genetics Applied to Livestock Production*.
- Khanna, A., Anumalla, M., Catolos, M., Bartholomé, J., Fritsche-Neto, R., Platten, J. D., Pisano, D. J., Gulles, A., Ma, M. T., Ramos, J., Faustino, G., Bhosale, S., & Hussain, W. (2022). Genetic Trends Estimation in IRRIs Rice Drought Breeding Program and Identification of High Yielding Drought-Tolerant Lines. *Rice*, 15(1). <https://doi.org/10.1186/s12284-022-00559-3>
- Kinghorn, B.P., Van Eenennaam, A.L., MacNeil, M.D. 2015. Targeting Multiple End-Uses in Beef Cattle Breeding Programs. *Proceedings of the 2015 Applied Reproductive Strategies in Beef Cattle conference*, August 17-18, 2015, Davis, CA.
- Knap, P. W. (2005). Breeding robust pigs. *Australian Journal of Experimental Agriculture*, 45(7–8), 763–773. <https://doi.org/10.1071/EA05041>
- Knap, P. W., & Su, G. (2008). Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. *Animal*, 2(12). <https://doi.org/10.1017/S1751731108003145>
- Knight, R. (1970). The measurement and interpretation of genotype-environment interactions. *Euphytica*, 19(2). <https://doi.org/10.1007/BF01902950>
- Kolmodin, R., Strandberg, E., Jorjani, H., & Danell, B. (2003). Selection in the presence of a genotype by environment interaction: response in environmental sensitivity. *Animal Science*, 76(3), 375–385. <https://doi.org/10.1017/s1357729800058604>
- Kolmodin, R., Strandberg, E., Madsen, P., Jensen, J., & Jorjani, H. (2002). Genotype by environment interaction in Nordic dairy cattle studied using reaction norms. *Acta Agriculturae Scandinavica - Section A: Animal Science*, 52(1). <https://doi.org/10.1080/09064700252806380>

- Kraakman, A. T. W., Niks, R. E., van den Berg, P. M. M. M., Stam, P., & van Eeuwijk, F. A. (2004). Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics*, *168*(1). <https://doi.org/10.1534/genetics.104.026831>
- Kusmec, A., Srinivasan, S., Nettleton, D., & Schnable, P. S. (2017). Distinct genetic architectures for phenotype means and plasticities in *Zea mays*. *Nature Plants*, *3*(9), 715–723. <https://doi.org/10.1038/s41477-017-0007-7>
- Lacaze, X., Hayes, P. M., & Korol, A. (2009). Genetics of phenotypic plasticity: QTL analysis in barley, *Hordeum vulgare*. *Heredity*, *102*(2), 163–173. <https://doi.org/10.1038/hdy.2008.76>
- Lee, S. H., & van der Werf, J. H. J. (2016). MTG2: An efficient algorithm for multivariate linear mixed model analysis based on genomic information. *Bioinformatics*, *32*(9), 1420–1422. <https://doi.org/10.1093/bioinformatics/btw012>
- Li, C., Wu, X., Li, Y., Shi, Y., Song, Y., Zhang, D., Li, Y., & Wang, T. (2019). Genetic architecture of phenotypic means and plasticities of kernel size and weight in maize. *Theoretical and Applied Genetics*, *132*(12), 3309–3320. <https://doi.org/10.1007/s00122-019-03426-w>
- Li, L., & Hermes, S. (2013). Genotype By Environment Interactions for Average Daily Gain Using Multiple-Trait Analyses in Australian Pigs. *Proc. Assoc. Advmt. Anim. Breed. Genet.*, *20*.
- Li, L., & Hermes, S. (2016). Evaluation of sire by environment interactions for growth rate and backfat depth using reaction norm models in pigs. *Journal of Animal Breeding and Genetics*, *133*(5). <https://doi.org/10.1111/jbg.12207>
- Li, L., Swan, A. A., Brown, D. J., & van der Werf, J. H. J. (2015). Australian Sheep Breeding Values for Worm Egg Count Retain Predictive Power Across Flocks in the Presence of GxE. *Proc. Assoc. Advmt. Breed. Genet.*, *386–389*.
- Li, L., Swan, A. A., Brown, D. J., & van der Werf, J. H. J. (2019). Genotype by environment interaction in Australian maternal and terminal sheep. *Proc. Assoc. Advmt. Breed. Genet.*, *234–237*.
- Li, X., Guo, T., Mu, Q., Li, X., & Yu, J. (2018). Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proceedings of the National Academy of*

Sciences of the United States of America, 115(26).
<https://doi.org/10.1073/pnas.1718326115>

- Lillehammer, M., Ødegård, J., & Meuwissen, T. H. (2009). Reducing the bias of estimates of genotype by environment interactions in random regression sire models. *Genetics Selection Evolution*, 41(1). <https://doi.org/10.1186/1297-9686-41-30>
- Lisle, C., Smith, A., Birrell, C. L., & Cullis, B. (2021). Information Based Diagnostic for Genetic Variance Parameter Estimation in Multi-Environment Trials. *Frontiers in Plant Science*, 12. <https://doi.org/10.3389/fpls.2021.785430>
- Lopez-Cruz, M., Crossa, J., Bonnett, D., Dreisigacker, S., Poland, J., Jannink, J. L., Singh, R. P., Autrique, E., & de los Campos, G. (2015). Increased prediction accuracy in wheat breeding trials using a marker $\times$ environment interaction genomic selection model. *G3: Genes, Genomes, Genetics*, 5(4). <https://doi.org/10.1534/g3.114.016097>
- Ly, D., Huet, S., Gauffreteau, A., Rincet, R., Touzy, G., Mini, A., Jannink, J. L., Cormier, F., Paux, E., Lafarge, S., Le Gouis, J., & Charmet, G. (2018). Whole-genome prediction of reaction norms to environmental stress in bread wheat (*Triticum aestivum* L.) by genomic random regression. *Field Crops Research*, 216, 32–41. <https://doi.org/10.1016/j.fcr.2017.08.020>
- MacLeod, I. M., Bowman, P. J., Vander Jagt, C. J., Haile-Mariam, M., Kemper, K. E., Chamberlain, A. J., Schrooten, C., Hayes, B. J., & Goddard, M. E. (2016). Exploiting biological priors and sequence variants enhances QTL discovery and genomic prediction of complex traits. *BMC Genomics*, 17(144), 1–21. <https://doi.org/10.1186/s12864-016-2443-6>
- Madsen, M. D., Madsen, P., Nielsen, B., Kristensen, T. N., Jensen, J., & Shirali, M. (2018). Macro-environmental sensitivity for growth rate in Danish Duroc pigs is under genetic control. *Journal of Animal Science*, 96(12). <https://doi.org/10.1093/jas/sky376>
- Mahjourimajd, S., Kuchel, H., Langridge, P., & Okamoto, M. (2016). Evaluation of Australian wheat genotypes for response to variable nitrogen application. *Plant and Soil*, 399(1–2). <https://doi.org/10.1007/s11104-015-2694-z>

- Malosetti, M., Ribaut, J. M., & van Eeuwijk, F. A. (2013). The statistical analysis of multi-environment data: Modeling genotype-by-environment interaction and its genetic basis. *Frontiers in Physiology*, 4 MAR. <https://doi.org/10.3389/fphys.2013.00044>
- Mart, K. B., Veneklaas, E. J., & Bramley, H. (2016). Osmotic potential at full turgor: An easily measurable trait to help breeders select for drought tolerance in wheat. *Plant Breeding*, 135(3). <https://doi.org/10.1111/pbr.12367>
- Meuwissen, T. H. E., De Jong, G., & Engel, B. (1996). Joint Estimation of Breeding Values and Heterogeneous Variances of Large Data Files. *Journal of Dairy Science*, 79(2). [https://doi.org/10.3168/jds.S0022-0302\(96\)76365-8](https://doi.org/10.3168/jds.S0022-0302(96)76365-8)
- Meyer, K. (2005). Random regression analyses using B-splines to model growth of Australian Angus cattle. *Genetics Selection Evolution*, 37(6), 473. <https://doi.org/10.1186/1297-9686-37-6-473>
- Meyer, K. (2009). Factor-analytic models for genotype $\times$ environment type problems and structured covariance matrices. *Genetics Selection Evolution*, 41(1). <https://doi.org/10.1186/1297-9686-41-21>
- Mickelbart, M. V., Hasegawa, P. M., & Bailey-Serres, J. (2015). Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. In *Nature Reviews Genetics* 16(4). <https://doi.org/10.1038/nrg3901>
- Millet, E. J., Kruijer, W., Coupel-Ledru, A., Alvarez Prado, S., Cabrera-Bosquet, L., Lacube, S., Charcosset, A., Welcker, C., van Eeuwijk, F., & Tardieu, F. (2019). Genomic prediction of maize yield across European environmental conditions. In *Nature Genetics* 51(6). <https://doi.org/10.1038/s41588-019-0414-y>
- Monaghan, P. (2008). Early growth conditions, phenotypic development and environmental change. In *Philosophical Transactions of the Royal Society B: Biological Sciences* 363(1497). <https://doi.org/10.1098/rstb.2007.0011>
- Mortimer, S. I., Hatcher, S., Fogarty, N. M., van der Werf, J. H. J., Brown, D. J., Swan, A. A., Greeff, J. C., Refshauge, G., Hocking Edwards, J. E., & Gaunt, G. M. (2017). Genetic parameters for wool traits, live weight, and ultrasound carcass traits in merino sheep. *Journal of Animal Science*, 95(5), 1879–1891. <https://doi.org/10.2527/jas2016.1234>

- Mota, L. F. M., Fernandes, G. A., Herrera, A. C., Scaletz, D. C. B., Espigolan, R., Magalhães, A. F. B., Carvalheiro, R., Baldi, F., & Albuquerque, L. G. (2020). Genomic reaction norm models exploiting genotype $\times$ environment interaction on sexual precocity indicator traits in Nellore cattle. *Animal Genetics*, 51(2). <https://doi.org/10.1111/age.12902>
- Mulder, H. A. (2016). Genomic selection improves response to selection in resilience by exploiting genotype by environment interactions. *Frontiers in Genetics*, 7(OCT). <https://doi.org/10.3389/fgene.2016.00178>
- Mulder, H. A. (2017). Is GxE a burden or a blessing? Opportunities for genomic selection and big data. *Journal of Animal Breeding and Genetics*, 134(6), 435–436. <https://doi.org/https://doi.org/10.1111/jbg.12303>
- Mulder, H. A., & Bijma, P. (2005). Effects of genotype x environment interaction on genetic gain in breeding programs. *Journal of Animal Science*, 83(1), 49–61. <https://doi.org/10.2527/2005.83149x>
- Mulder, H. A., Groen, A. F., De Jong, G., & Bijma, P. (2004). Genotype x environment interaction for yield and somatic cell score with automatic and conventional milking systems. *Journal of Dairy Science*, 87(5). [https://doi.org/10.3168/jds.S0022-0302\(04\)73300-7](https://doi.org/10.3168/jds.S0022-0302(04)73300-7)
- Mulder, H. A., Rönnegård, L., Fikse, W. F., Veerkamp, R. F., & Strandberg, E. (2013). Estimation of genetic variance for macro- and micro-environmental sensitivity using double hierarchical generalized linear models. *Genetics Selection Evolution*, 45(1). <https://doi.org/10.1186/1297-9686-45-23>
- Mulder, H. A., Veerkamp, R. F., Ducro, B. J., Van Arendonk, J. A. M., & Bijma, P. (2006). Optimization of dairy cattle breeding programs for different environments with genotype by environment interaction. *Journal of Dairy Science*, 89(5). [https://doi.org/10.3168/jds.S0022-0302\(06\)72242-1](https://doi.org/10.3168/jds.S0022-0302(06)72242-1)
- Nabugoomu, F., Kempton, R. A., & Talbot, M. (1999). Analysis of series of trials where varieties differ in sensitivity to locations. *Journal of Agricultural, Biological, and Environmental Statistics*, 4(3). <https://doi.org/10.2307/1400388>
- Namkoong, G. (1985). The influence of composite traits on genotype by environment relations. *Theoretical and Applied Genetics*, 70(3). <https://doi.org/10.1007/BF00304918>

- Nel, C. L., Steyn, Y., Gilmour, A. R., Waters, D., Clark, S. A., van der Werf, J. H. J., Swan, A. A., Dzama, K., & Cloete, S. W. P. (2023). Reaction-norm analysis of neonatal lamb mortality suggests heritability varies with cold-stress: an example in the Elsenburg Merino selection lines. *Animal Production Science*. <https://doi.org/10.1071/AN22464>
- Nguyen, T. T. T., Bowman, P. J., Haile-Mariam, M., Nieuwhof, G. J., Hayes, B. J., & Pryce, J. E. (2017). Short communication: Implementation of a breeding value for heat tolerance in Australian dairy cattle. *Journal of Dairy Science*, 100(9). <https://doi.org/10.3168/jds.2017-12898>
- Nguyen, T. T. T., Bowman, P. J., Haile-Mariam, M., Pryce, J. E., & Hayes, B. J. (2016). Genomic selection for tolerance to heat stress in Australian dairy cattle. *Journal of Dairy Science*, 99(4). <https://doi.org/10.3168/jds.2015-9685>
- Ni, G., van der Werf, J., Zhou, X., Hyppönen, E., Wray, N. R., & Lee, S. H. (2019). Genotype–covariate correlation and interaction disentangled by a whole-genome multivariate reaction norm model. *Nature Communications*, 10(1). <https://doi.org/10.1038/s41467-019-10128-w>
- Nicholls, N., Drosowsky, W., & Lavery, B. (1997). Australian rainfall variability and change. *Weather*, 52(3), 66–72. <https://doi.org/10.1002/j.1477-8696.1997.tb06274.x>
- Nilforooshan, M. A., & Jorjani, H. (2022). Invited review: A quarter of a century—International genetic evaluation of dairy sires using MACE methodology. In *Journal of Dairy Science*, 105(1). <https://doi.org/10.3168/jds.2021-20927>
- Obst, J. M., & Day, H. R. (1968). The effect of inclement weather on mortality of merino and corriedale lambs on kangaroo island. *Proceedings of the Australian Society of Animal Production*.
- Oldham, C. M., Thompson, A. N., Ferguson, M. B., Gordon, D. J., Kearney, G. A., & Paganoni, B. L. (2011). The birthweight and survival of Merino lambs can be predicted from the profile of liveweight change of their mothers during pregnancy. *Animal Production Science*, 51(9). <https://doi.org/10.1071/AN10155>
- Oliveira, D. P., Lourenco, D. A. L., Tsuruta, S., Misztal, I., Santos, D. J. A., de Araújo Neto, F. R., Aspilcueta-Borquis, R. R., Baldi, F., Carnevalheiro, R., de Camargo, G. M. F., Albuquerque, L. G., & Tonhati, H. (2018). Reaction norm for yearling weight in beef

- cattle using single-step genomic evaluation. *Journal of Animal Science*, 96(1), 27–34. <https://doi.org/10.1093/jas/skx006>
- Oliveira, D. P., Lourenco, D. A. L., Tsuruta, S., Misztal, I., Santos, D. J. A., De Araújo Neto, F. R., Aspilcueta-Borquis, R. R., Baldi, F., Carneiro, R., De Camargo, G. M. F., Albuquerque, L. G., & Tonhati, H. (2018). Reaction norm for yearling weight in beef cattle using single-step genomic evaluation. *Journal of Animal Science*, 96(1), 27–34. <https://doi.org/10.1093/jas/skx006>
- Olivieri, B. F., Mercadante, M. E. Z., Cyrillo, J. N. D. S. G., Branco, R. H., Bonilha, S. F. M., De Albuquerque, L. G., De Oliveira Silva, R. M., & Baldi, F. (2016). Genomic regions associated with feed efficiency indicator traits in an experimental nelore cattle population. *PLoS ONE*, 11(10), 1–19. <https://doi.org/10.1371/journal.pone.0164390>
- Paschke, M., Bernasconi, G., & Schmid, B. (2003). Population Size and Identity Influence the Reaction Norm of the Rare, Endemic Plant *Cochlearia bavarica* across a Gradient of Environmental Stress. *Evolution*, 57(3), 496–508. <https://doi.org/10.1111/j.0014-3820.2003.tb01541.x>
- Perkins, J. M., & Jinks, J. L. (1968). Environmental and genotype-environmental components of variability III. Multiple lines and crosses. *Heredity*, 23(3). <https://doi.org/10.1038/hdy.1968.48>
- Pollot, G. E., & Greeff, J. C. (2004). Genotype × environment interactions and genetic parameters for fecal egg count and production traits of Merino sheep. *Journal of Animal Science*, 82, 2840–2851.
- Pollott, G. E., & Greeff, J. C. (2004). Genotype x environment interactions and genetic parameters for fecal egg count and production traits of Merino sheep. *Journal of Animal Science*, 82(10). <https://doi.org/10.2527/2004.82102840x>
- Powell, N., Ji, X., Ravash, R., Edlington, J., & Dolferus, R. (2012). Yield stability for cereals in a changing climate. In *Functional Plant Biology*, (39)7. <https://doi.org/10.1071/FP12078>
- Qian, L., Hickey, L. T., Stahl, A., Werner, C. R., Hayes, B., Snowdon, R. J., & Voss-Fels, K. P. (2017). Exploring and harnessing haplotype diversity to improve yield stability in crops. *Frontiers in Plant Science*, 8. <https://doi.org/10.3389/fpls.2017.01534>

- Raffo, M. A., Sarup, P., Andersen, J. R., Orabi, J., Jahoor, A., & Jensen, J. (2022). Integrating a growth degree-days based reaction norm methodology and multi-trait modeling for genomic prediction in wheat. *Frontiers in Plant Science*, 13. <https://doi.org/10.3389/fpls.2022.939448>
- Raman, H., Shamaya, N., Pirathiban, R., McVittie, B., Raman, R., Cullis, B., & Easton, A. (2023). Quantitative Trait Loci for Genotype and Genotype by Environment Interaction Effects for Seed Yield Plasticity to Terminal Water-Deficit Conditions in Canola (*Brassica napus* L.). *Plants*, 12(4). <https://doi.org/10.3390/plants12040720>
- Rauw, W. M., & Gomez-Raya, L. (2015). Genotype by environment interaction and breeding for robustness in livestock. In *Frontiers in Genetics*, 6 <https://doi.org/10.3389/fgene.2015.00310>
- Robertson, A. (1959). The Sampling Variance of the Genetic Correlation Coefficient. *Biometrics*, 15(3). <https://doi.org/10.2307/2527750>
- Rojas-Downing, M. M., Nejadhashemi, A. P., Harrigan, T., & Woznicki, S. A. (2017). Climate change and livestock: Impacts, adaptation, and mitigation. In *Climate Risk Management*, 6 <https://doi.org/10.1016/j.crm.2017.02.001>
- Saavedra-Jiménez, L. A., Ramírez-Valverde, R., Núñez-Domínguez, R., García-Muñiz, J. G., Lopez-Villalobos, N., & Ruíz-Flores, A. (2013). Genotype by climate interaction in the genetic evaluation for growing traits of Braunvieh cattle in Mexico. *Tropical Animal Health and Production*, 45(7). <https://doi.org/10.1007/s11250-013-0387-2>
- Sadras, V. O., Montoro, A., Moran, M. A., & Aphalo, P. J. (2012). Elevated temperature altered the reaction norms of stomatal conductance in field-grown grapevine. *Agricultural and Forest Meteorology*, 165. <https://doi.org/10.1016/j.agrformet.2012.06.005>
- Sae-Lim, P., Gjerde, B., Nielsen, H. M., Mulder, H., & Kause, A. (2016). A review of genotype-by-environment interaction and micro-environmental sensitivity in aquaculture species. In *Reviews in Aquaculture* 8(4). <https://doi.org/10.1111/raq.12098>
- Sae-Lim, P., Komen, H., Kause, A., & Mulder, H. A. (2014). Identifying environmental variables explaining genotype-by-environment interaction for body weight of rainbow trout (*Onchorynchus mykiss*): Reaction norm and factor analytic models. *Genetics Selection Evolution*, 46(1). <https://doi.org/10.1186/1297-9686-46-16>

- Sae-Lim, P., Mulder, H., Gjerde, B., Koskinen, H., Lillehammer, M., & Kause, A. (2015). Genetics of growth reaction norms in farmed rainbow trout. *PLoS ONE*, *10*(8). <https://doi.org/10.1371/journal.pone.0135133>
- Sandhu, K. S., Mihalyov, P. D., Lewien, M. J., Pumphrey, M. O., & Carter, A. H. (2021). Genomic Selection and Genome-Wide Association Studies for Grain Protein Content Stability in a Nested Association Mapping Population of Wheat. *Agronomy*, *11*(12). <https://doi.org/10.3390/agronomy11122528>
- Santana, M. L., Bignardi, A. B., Eler, J. P., Cardoso, F. F., & Ferraz, J. B. S. (2013). Genotype by environment interaction and model comparison for growth traits of Santa Ines sheep. *Journal of Animal Breeding and Genetics*, *130*(5). <https://doi.org/10.1111/jbg.12029>
- Santana, M. L., Eler, J. P., Bignardi, A. B., Menéndez-Buxadera, A., Cardoso, F. F., & Ferraz, J. B. S. (2015). Multi-trait linear reaction norm model to describe the pattern of phenotypic expression of some economic traits in beef cattle across a range of environments. *Journal of Applied Genetics*, *56*(2). <https://doi.org/10.1007/s13353-014-0242-9>
- Seixas, L., De Melo, C. B., Tanure, C. B., Peripolli, V., & McManus, C. (2017). Heat tolerance in Brazilian hair sheep. *Asian-Australasian Journal of Animal Sciences*, *30*(4). <https://doi.org/10.5713/ajas.16.0191>
- Sheep Genetics. (2022). *Annual Outcomes Report 2021-21*.
- Shi, R., Brito, L. F., Liu, A., Luo, H., Chen, Z., Liu, L., Guo, G., Mulder, H., Ducro, B., van der Linden, A., & Wang, Y. (2021). Genotype-by-environment interaction in Holstein heifer fertility traits using single-step genomic reaction norm models. *BMC Genomics*, *22*(1). <https://doi.org/10.1186/s12864-021-07496-3>
- Silva, F. F., Guimarães, S. E. F., Lopes, P. S., Mulder, H. A., Bastiaansen, J. W. M., Knol, E. F., Lopes, M. S., Mathur, P. K., & Viana, J. M. S. (2014). Sire evaluation for total number born in pigs using a genomic reaction norms approach. *Journal of Animal Science*, *92*(9). <https://doi.org/10.2527/jas.2013-6486>
- Silva, F. F., Mulder, H. A., Knol, E. F., & Lopes, M. S. (2014). Sire evaluation for total number born in pigs using a genomic reaction norms approach. *Journal of Animal Science*, *92*, 3825–3834. <https://doi.org/10.2527/jas2013-6486>

- Sjoberg, S. M., Carter, A. H., Steber, C. M., & Garland Campbell, K. A. (2021). Application of the factor analytic model to assess wheat falling number performance and stability in multienvironment trials. *Crop Science*, 61(1). <https://doi.org/10.1002/csc2.20293>
- Sjoberg, S. M., Carter, A. H., Steber, C. M., & Garland-Campbell, K. A. (2020). Unraveling complex traits in wheat: Approaches for analyzing genotype $\times$ environment interactions in a multienvironment study of falling numbers. *Crop Science*, 60(6), 3013–3026. <https://doi.org/https://doi.org/10.1002/csc2.20133>
- Smith, A. B., Cullis, B., & Gilmour, A. (2001). Applications: The Analysis of Crop Variety Evaluation Data in Australia. *Australian & New Zealand Journal of Statistics*, 43(2), 129–145. <https://doi.org/https://doi.org/10.1111/1467-842X.00163>
- Smith, A. B., & Cullis, B. R. (2018). Plant breeding selection tools built on factor analytic mixed models for multi-environment trial data. *Euphytica*, 214(8). <https://doi.org/10.1007/s10681-018-2220-5>
- Smith, A. B., Cullis, B. R., & Thompson, R. (2005). The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. In *Journal of Agricultural Science* 143(6). <https://doi.org/10.1017/S0021859605005587>
- Smith, A. B., Ganesalingam, A., Kuchel, H., & Cullis, B. R. (2015). Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *TAG. Theoretical and Applied Genetics*, 128(1). <https://doi.org/10.1007/s00122-014-2412-x>
- Smith, A. B., Norman, A., Kuchel, H., & Cullis, B. (2021). Plant Variety Selection Using Interaction Classes Derived From Factor Analytic Linear Mixed Models: Models With Independent Variety Effects. *Frontiers in Plant Science*, 12. <https://doi.org/10.3389/fpls.2021.737462>
- Stefani, G., Santana Júnior, M. L., El Faro, L., & Tonhati, H. (2022). Genetics of tolerance to heat stress in milk yield of dairy buffaloes assessed by a reaction norm model. *Journal of Animal Breeding and Genetics*, 139(2). <https://doi.org/10.1111/jbg.12654>
- Strandén, I., & Garrick, D. J. (2009). Technical note: Derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. *Journal of Dairy Science*, 92(6), 2971–2975. <https://doi.org/10.3168/jds.2008-1929>

- Strillacci, M. G., Frigo, E., Schiavini, F., Samoré, A. B., Canavesi, F., Vevey, M., Cozzi, M. C., Soller, M., Lipkin, E., & Bagnato, A. (2014). Genome-wide association study for somatic cell score in Valdostana Red Pied cattle breed using pooled DNA. *BMC Genetics*, 15(1). <https://doi.org/10.1186/s12863-014-0106-7>
- Su, G., Madsen, P., Lund, M. S., Sorensen, D., Korsgaard, I. R., & Jensen, J. (2006). Bayesian analysis of the linear reaction norm model with unknown covariates. *Journal of Animal Science*, 84(7). <https://doi.org/10.2527/jas.2005-517>
- Tan, Y., Zhou, J., Wang, J., & Sun, L. (2020). The Genetic Architecture for Phenotypic Plasticity of the Rice Grain Ionome. *Frontiers in Plant Science*, 11. <https://doi.org/10.3389/fpls.2020.00012>
- Telfer, P., Edwards, J., Norman, A., Bennett, D., Smith, A., Able, J. A., & Kuchel, H. (2021). Genetic analysis of wheat (*Triticum aestivum*) adaptation to heat stress. *Theoretical and Applied Genetics*, 134(5), 1387–1407. <https://doi.org/10.1007/s00122-021-03778-2>
- Telfer, P., Edwards, J., Taylor, J., Able, J. A., & Kuchel, H. (2022). A multi-environment framework to evaluate the adaptation of wheat (*Triticum aestivum*) to heat stress. *Theoretical and Applied Genetics*, 135(4). <https://doi.org/10.1007/s00122-021-04024-5>
- Tolhurst, D. J., Gaynor, R. C., Gardunia, B., Hickey, J. M., & Gorjanc, G. (2022). Genomic selection using random regressions on known and latent environmental covariates. *Theoretical and Applied Genetics*. <https://doi.org/10.1007/s00122-022-04186-w>
- Tolhurst, D. J., Mathews, K. L., Smith, A. B., & Cullis, B. R. (2019). Genomic selection in multi-environment plant breeding trials using a factor analytic linear mixed model. *Journal of Animal Breeding and Genetics*, 136(4). <https://doi.org/10.1111/jbg.12404>
- Troyer, A. F. (1996). Breeding widely adapted, popular maize hybrids. *Euphytica*, 92(1-2), 163–174. <https://doi.org/10.1007/BF00022842>
- Tsuruta, S., Lourenco, D. A. L., Misztal, I., & Lawlor, T. J. (2015). Genotype by environment interactions on culling rates and 305-day milk yield of Holstein cows in 3 US regions. *Journal of Dairy Science*, 98(8), 5796–5805. <https://doi.org/10.3168/jds.2014-9242>
- Tyrisevä, A. M., Meyer, K., Fikse, W. F., Ducrocq, V., Jakobsen, J., Lidauer, M. H., & Mäntysaari, E. A. (2011). Principal component and factor analytic models in international

- sire evaluation. *Genetics Selection Evolution*, 43(1). <https://doi.org/10.1186/1297-9686-43-33>
- van der Putten, W., Poorter, H., Fiorani, F., Pieruschka, R., Wojciechowski, T., Schurr, U., Postma, J., van der Putten, W. H., & Kleyer, M. (2016). Pampered inside, pestered outside? Differences and similarities between plants growing in controlled conditions and in the field. *New Phytologist*, 212(4).
- van der Werf, J., & Banks, R. (2017). Effective breeding programmes down under. In *Journal of Animal Breeding and Genetics* 134(1). <https://doi.org/10.1111/jbg.12252>
- van Der Werf, J. H. J. (2004). Is it useful to define residual feed intake as a trait in animal breeding programs? *Australian Journal of Experimental Agriculture*, 44(4–5), 405–409. <https://doi.org/10.1071/ea02105>
- van der Werf, J. H. J., Banks, R. G., & Dekkers, J. C. M. (2018). Optimizing multiple trait selection. *Proceedings of the World Congress on Genetics Applied to Livestock Production*, 11, 889.
- van der Werf, J. H. J., Kinghorn, B. P., & Banks, R. G. (2010). Design and role of an information nucleus in sheep breeding programs. *Animal Production Science*, 50(11–12), 998–1003. <https://doi.org/10.1071/AN10151>
- van Der Werf, J. H. J., Kinghorn, B. P., & Banks, R. G. (2010). Design and role of an information nucleus in sheep breeding programs. *Animal Production Science*, 50(11–12). <https://doi.org/10.1071/AN10151>
- van Eeuwijk, F. A., Bustos-Korts, D. V., & Malosetti, M. (2016). What should students in plant breeding know about the statistical aspects of genotype $\times$ Environment interactions? *Crop Science*, 56(5). <https://doi.org/10.2135/cropsci2015.06.0375>
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91(11), 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Vanvanhossou, S. F. U., Scheper, C., Dossa, L. H., Yin, T., Brügemann, K., & König, S. (2020). A multi-breed GWAS for morphometric traits in four Beninese indigenous cattle breeds reveals loci associated with conformation, carcass and adaptive traits. *BMC Genomics*, 21(1), 1–16. <https://doi.org/10.1186/s12864-020-07170-0>

- Vercoe, J. E., & Frisch, J. E. (1992). Genotype (breed) and environment interaction with particular reference to cattle in the tropics - Review -. *Asian-Australasian Journal of Animal Sciences*, 5(3). <https://doi.org/10.5713/ajas.1992.401>
- Via, S., Gomulkiewicz, R., De Jong, G., Scheiner, S. M., Schlichting, C. D., & Van Tienderen, P. H. (1995). Adaptive phenotypic plasticity: consensus and controversy. In *Trends in Ecology & Evolution*, 10(5). [https://doi.org/10.1016/S0169-5347\(00\)89061-8](https://doi.org/10.1016/S0169-5347(00)89061-8)
- Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era - Concepts and misconceptions. In *Nature Reviews Genetics* 9(4). <https://doi.org/10.1038/nrg2322>
- Wahinya, P. K., Jeyaruban, G., Swan, A., & Magothe, T. (2020). Estimation of genetic parameters for milk and fertility traits within and between low, medium and high dairy production systems in Kenya to account for genotype-by-environment interaction. *Journal of Animal Breeding and Genetics*, 137(5). <https://doi.org/10.1111/jbg.12473>
- Wallenbeck, A., Rydhmer, L., & Lundeheim, N. (2009). GxE interactions for growth and carcass leanness: Re-ranking of boars in organic and conventional pig production. *Livestock Science*, 123(2–3). <https://doi.org/10.1016/j.livsci.2008.11.003>
- Wang, Y., Saelao, P., Kern, C., Jin, S., Gallardo, R. A., Kelly, T., Dekkers, J. M., Lamont, S. J., & Zhou, H. (2020). Liver transcriptome responses to heat stress and newcastle disease virus infection in genetically distinct chicken inbred lines. *Genes*, 11(9), 1–30. <https://doi.org/10.3390/genes11091067>
- Waters, D. L., Clark, S. A., Moghaddar, N., & Van der Werf, J. H. (2022). Genomic analysis of the slope of the reaction norm for body weight in Australian sheep. *Genetics Selection Evolution*, 54(1). <https://doi.org/10.1186/s12711-022-00734-6>
- Waters, D. L., Van Der Werf, J. H. J., & Clark, S. A. (2022). Building on Falconer’s work on environmental sensitivity and scale. *Proceedings of the 12th World Congress on Genetics Applied to Livestock Production*.
- Waters, D. L., van der Werf, J. H. J., Robinson, H., Hickey, L. T., & Clark, S. A. (2023). Partitioning the forms of genotype-by-environment interaction in the reaction norm analysis of stability. *Theoretical and Applied Genetics*, 136(5), 99. <https://doi.org/10.1007/s00122-023-04319-9>

- Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M.-D., Asyraf Md Hatta, M., Hinchliffe, A., Steed, A., Reynolds, D., Adamski, N. M., Breakspear, A., Korolev, A., Rayner, T., Dixon, L. E., Riaz, A., Martin, W., Ryan, M., Edwards, D., ... Hickey, L. T. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nature Plants*, 4(1), 23–29. <https://doi.org/10.1038/s41477-017-0083-8>
- Wellmann, R. (2023). Selection index theory for populations under directional and stabilizing selection. *Genetics Selection Evolution*, 55(10). <https://doi.org/10.1186/s12711-023-00776-4>
- West, J. W. (2003). Effects of heat-stress on production in dairy cattle. *Journal of Dairy Science*, 86(6). [https://doi.org/10.3168/jds.S0022-0302\(03\)73803-X](https://doi.org/10.3168/jds.S0022-0302(03)73803-X)
- Westell, R. A., Quaas, R. L., & Van Vleck, L. D. (1988). Genetic Groups in an Animal Model. *Journal of Dairy Science*, 71(5), 1310–1318. [https://doi.org/10.3168/jds.S0022-0302\(88\)79688-5](https://doi.org/10.3168/jds.S0022-0302(88)79688-5)
- Whelan, M. B., Cottle, D. J., Geenty, K. G., & Brown, D. J. (2009). Classifying sheep grazing environments using satellite data to quantify genotype by environment interactions. *Proc. Assoc. Advmt. Anim. Breed. Genet*, 52–55.
- Wiggans, G. R., & VanRaden, P. M. (1991). Method and Effect of Adjustment for Heterogeneous Variance. *Journal of Dairy Science*, 74(12). [https://doi.org/10.3168/jds.S0022-0302\(91\)78631-1](https://doi.org/10.3168/jds.S0022-0302(91)78631-1)
- Williams, J. L., Bertrand, J. K., Misztal, I., & Łukaszewicz, M. (2012). Genotype by environment interaction for growth due to altitude in United States Angus cattle. *Journal of Animal Science*, 90(7). <https://doi.org/10.2527/jas.2011-4365>
- Woltereck, R. (1913). Weitere experimentelle untersuchungen über Artänderung, speziell über das Wesen quantitativer Artunterschiede bei Daphniden. *Zeitschrift Für Induktive Abstammungs- Und Vererbungslehre*, 9(1), 146. <https://doi.org/10.1007/BF01876686>
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., & Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42(7). <https://doi.org/10.1038/ng.608>

- Yates, F., & Cochran, W. G. (1938a). The analysis of groups of experiments. *The Journal of Agricultural Science*, 28(4). <https://doi.org/10.1017/S0021859600050978>
- Yates, F., & Cochran, W. G. (1938b). The analysis of groups of experiments. *The Journal of Agricultural Science*, 28(4). <https://doi.org/10.1017/S0021859600050978>
- Zeven, A. C. (1998). Landraces: A review of definitions and classifications. *Euphytica*, 104(2). <https://doi.org/10.1023/A:1018683119237>
- Zhang, Z., Kargo, M., Liu, A., Thomasen, J. R., Pan, Y., & Su, G. (2019). Genotype-by-environment interaction of fertility traits in Danish Holstein cattle using a single-step genomic reaction norm model. *Heredity*, 123, 202–214. <https://doi.org/10.1038/s41437-019-0192-4>
- Zhao, R., Muehlbauer, E., Decuypere, E., & Grossmann, R. (2004). Effect of genotype-nutrition interaction on growth and somatotrophic gene expression in the chicken. *General and Comparative Endocrinology*, 136(1). <https://doi.org/10.1016/j.ygcen.2003.11.009>