

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

ESTIMATION OF GENETIC AND PHENOTYPIC  
PARAMETERS IN NEW ZEALAND ROMNEY SHEEP

A thesis presented in partial fulfilment of  
the requirements for the degree of  
Masterate of Agricultural Science  
in Animal Science  
at Massey University

SUSAN JANE TAIT

1983

ABSTRACT

Genetic and phenotypic parameters were estimated from liveweight and fleece data recorded on 1604 New Zealand Romney lambs between 1970 and 1972. The flock into which the lambs were born is located at Woodlands Research Station near Invercargill. The data analysed are from the establishment phase of a long-term selection experiment; the flock was closed for selection in 1973.

The traits studied were birthweight (BWT), docking weight (DWT), weaning weight (WWT), April liveweight (APR), June liveweight (JUN), August liveweight (AUG), November liveweight (NOV), 2-tooth liveweight (2TH), lamb fleece weight (LFW), hogget fleece weight (HFW), staple length (STL), quality number (QNO), character (CHR), fleece colour (COL) and break severity (BRS).

Restricted maximum likelihood (REML) estimates of the variance components were obtained. These were used in the generation of paternal half-sib estimates of the heritabilities ( $h^2$ ), the inter-trait genetic ( $r_g$ ) and phenotypic ( $r_p$ ) correlations, and the best linear unbiased estimates (BLUE) of the non-genetic (fixed) effects.

The estimates of the  $h^2$ 's for the liveweights ranged from 0.08 for BWT, increasing through to 0.13 for 2TH. These estimates are lower than most of the values previously published (generally, from 0.2-0.4, respectively), although they are comparable with many of the more recent  $h^2$  estimates for liveweight.

The estimates of  $h^2$  for the fleece traits were generally similar to the estimates of previous studies. Estimates of 0.19 and 0.30 were obtained for LFW and HFW, respectively, and 0.37 for STL. The

fleece quality traits were found to have  $h^2$  estimates ranging from 0.07 for BRS to 0.56 for QNO.

The estimates of the genetic and phenotypic correlations between the traits studied were comparable with estimates from previous studies in most cases. Important exceptions include the low genetic correlations of WWT with the liveweights from JUN (of 0.38) through to 2TH (of 0.50).

The BLUE's of the fixed effects generally agree well with the estimates of previous studies. Year-of-birth, birth-rearing rank and date-of-birth effects were significant for all the traits studied. In addition, age-of-dam effects were significant for all the liveweights, and sex effects were significant for BWT, DWT and WWT.

ACKNOWLEDGEMENTS

I would first like to sincerely thank my supervisors, Professor R.D. Anderson and Professor A.L. Rae for the invaluable advice and assistance they have given throughout this study.

The efforts of the staff at the Genetics Section of Ruakura Animal Research Station in providing the data for study are most appreciated. Thanks are particularly due to Dr J.N. Clarke and Sharon Hickey.

The financial help through grants and scholarships awarded by the Massey University Scholarships Committee is gratefully acknowledged.

The assistance of Arthur Gilmour in computer programming matters and the use of his generalised linear program, REG, is acknowledged with gratitude.

Thanks are also extended to Robin Vining for the care and attention given to the typing of this thesis. This was a great help in its completion.

A very appreciative thank you is extended to each of my close friends and members of my family who gave me so much support, especially in the final stages.

To Brennon Wood, who helped in many ways from beginning to end, I give my special thanks.

TABLE OF CONTENTS

<u>Chapter</u>		<u>Page</u>
	ABSTRACT	ii
	ACKNOWLEDGEMENTS	iv
	LIST OF TABLES	viii
ONE	INTRODUCTION	1
TWO	REVIEW OF LITERATURE	
	I. VARIANCE COMPONENT ESTIMATION THEORY	4
	A. Introduction	4
	B. Henderson's Methods	6
	1. Henderson's Method I	7
	2. Henderson's Method II	8
	3. Henderson's Method III	8
	C. Maximum Likelihood	10
	D. Restricted Maximum Likelihood	16
	E. Other Estimation Methods	21
	F. Computing Algorithms	25
	1. Mathematical optimization	25
	2. Unconstrained optimization algorithms	27
	3. Use of the mixed model equations	31
	4. Modifications to apply constraints	33
	II. PARAMETER ESTIMATES FOR THE ROMNEY	36
	A. Non-Genetic Effects	36
	B. Heritabilities	42
	C. Correlations	57

<u>Chapter</u>		<u>Page</u>
THREE	SOURCE OF DATA	76
	I. RUAKURA SHEEP SELECTION STUDIES	76
	A. Introduction and Objectives	76
	B. Design of the Experiment	77
	1. The establishment phase	77
	2. The selection phase	78
	II. ESTABLISHMENT PHASE OF THE WOODLANDS ROMNEY FLOCKS	81
	A. The Sheep and their Environment	81
	B. The Data Collected	83
FOUR	STATISTICAL METHODS	86
	I. RESTRICTED MAXIMUM LIKELIHOOD ESTIMATION OF VARIANCE COMPONENTS	86
	A. The Model Used	86
	B. The Computing Algorithm Used	89
	C. Comparisons with the Method III and MIVQUE Estimates of the Variance Components	89
	II. ESTIMATION OF HERITABILITY	91
	III. ESTIMATION OF THE CORRELATIONS	93
	IV. ESTIMATION OF THE NON-GENETIC (FIXED) EFFECTS	95
	A. The Models Used	95
	B. The Computing Algorithm Used	97
FIVE	RESULTS	99
SIX	DISCUSSION	112
	I. NON-GENETIC EFFECTS	112
	II. HERITABILITY	118
	III. CORRELATIONS	123

<u>Chapter</u>		<u>Page</u>
SEVEN	CONCLUSION	128
	BIBLIOGRAPHY	130
	APPENDIX	145

LIST OF TABLES

<u>Table</u>		<u>Page</u>
2.1	Correction factors used by Sheeplan for dual-purpose sheep breeds	40
2.2	Estimates of environment effects on weaning weight, hogget liveweight and hogget fleece weight in Romney sheep	41
2.3	Heritability estimates for various liveweight and fleece characteristics in Romney sheep	45
2.4	Phenotypic correlation estimates between various liveweight and fleece characteristics in Romney sheep	61
2.5	Genetic correlation estimates between various liveweight and fleece characteristics in Romney sheep	69
3.1	Sheep selection experiments at Tokanui, Templeton and Woodlands	80
3.2	Age at measurement and the abbreviations used for the liveweight and fleece traits	85
5.1	Estimates of the heritabilities from different methods of variance component estimation	102
5.2	Estimates of the genetic (above diagonal) and phenotypic (below diagonal) correlations between the liveweight and fleece traits	103
5.3	Significance of the environmental effects on the liveweights	104
5.4	Significance of the environmental effects on the logarithm of the liveweights	105
5.5	Significance of the environmental effects on the fleece traits	106
5.6	Estimates of the environmental effects (kg) on the liveweights with both sexes combined	107
5.7	Estimates of the environmental effects (kg) on the liveweights for each sex separately	108
5.8	Estimates of the environmental effects (kg) on the fleece traits for each sex separately	110

CHAPTER ONE

INTRODUCTION

## CHAPTER ONE

INTRODUCTION

The retention of a selected group of individuals for breeding purposes is a fundamental aspect of any selection programme. While the actual methods used to select between individuals may vary, they all aim to identify individuals with the highest genetic merit. An effective selection policy combined with a complementary mating plan forms the basis of a breeding programme designed to maximize genetic gain in each generation of selection. These concepts were applied to animal breeding during the 1930's and 40's by Lush and Hazel following earlier work by Fisher, Wright and Haldane (see, for example, Lush, 1937; Hazel, 1943; Fisher, 1930; Wright, 1939; Haldane, 1946).

Both single-trait improvement and selection for several traits simultaneously (e.g., using selection indices) can be breeding objectives. This is particularly so with a dual-purpose breed such as the Romney, where an aggregate breeding value may be estimated as a basis for choosing between individuals. It is important, therefore, to know how certain productive traits can be expected to change in future generations as a result of selection on one or several other traits. These factors are dependent on such parameters as the heritabilities of the individual traits and the genetic and phenotypic correlations between traits. These parameters are all functions of variance components; thus the estimation of variance components is an integral and important aspect of animal improvement.

As with any field of estimation, many procedures for estimating variance components have been developed. Each method of estimation

has its own properties and applicability to a specific situation. Not all methods can be applied in animal breeding settings where, in particular, the data are usually unbalanced (i.e., there are unequal numbers of observations in the different subclasses), and the underlying model is mixed (i.e., includes both fixed and random effects).

Variance components have traditionally been estimated by analysis of variance (ANOVA) type methods such as the methods outlined by Henderson (1953). These estimates have then been used to estimate genetic and phenotypic parameters, such as heritabilities and correlations, for use in breed improvement plans. However, developments in computing facilities have been accompanied by the emergence of a number of new methods for estimating variance components which can be applied to mixed models with unbalanced data.

In this study it was decided to use Restricted Maximum Likelihood (REML) to estimate the variance components. REML shares with Maximum Likelihood (ML) some desirable features of the maximum likelihood approach, such as translation invariance as well as "having the intuitively appealing property of maximizing the likelihood function" (Mood, Graybill and Boes, 1974), which is always well-defined. However, like ML, REML is an iterative procedure, which precludes unbiasedness properties, and does require assumptions concerning the form of the distribution of the data. When estimating variance components from mixed models, REML is preferable to ML because it takes account of the fixed effects in the model whereas ML does not. Furthermore, in the case of balanced data, REML has been shown to coincide with ANOVA estimators (which are best, unbiased, quadratic estimators, BQUE, in the balanced data case) (R.D. Anderson, 1978).

20ms  
Having generated the REML estimates of the variance components, advantage was taken of the fact that under normality the first iterate of the REML estimator is equal to the MIVQUE (and the MINQUE under normality) to also obtain MIVQUE estimates of the variance components. Method III estimates were also estimated and, in fact, were used as the a priori starting values for the MIVQUE estimates.

Environmental effects have traditionally been estimated by least squares procedures. In this study, the REML estimates of the variance components were regarded as the known parameter values under normality; allowing best linear unbiased estimates of the environmental effects to be obtained. Hence, this provides another interesting comparison.

This study provides a basis for evaluating genetic and phenotypic parameter estimates of a range of liveweight and fleece traits using a new method of estimation, namely; REML.

CHAPTER TWO

REVIEW OF LITERATURE

## CHAPTER TWO

REVIEW OF LITERATUREI. VARIANCE COMPONENT ESTIMATION THEORYA. Introduction

Fisher (1918) first used the terms "variance" and "analysis of variance" and later, (Fisher, 1925), he formally outlined an analysis of variance (ANOVA) pertaining to the 1-way random-effects model with balanced data. The estimation of variance components using the ANOVA approach was, according to R.D. Anderson (1978), clarified by Tippet (1931), having not been explicitly presented by Fisher (1925).

In animal breeding settings, the data are rarely balanced. It was Cochran (1939) who proposed a method for estimating variance components from unbalanced data, which is analogous to Fisher's (1925) technique. A clear demonstration of the case of an ANOVA approach to estimating variance components for the 1-way random-effects model with unbalanced data was given by Winsor and Clarke (1940). The ANOVA method was then "extended to the multi-fold nested classification model by Ganguli (1941) and to the 2-way crossed classification random model with interaction by Crump (1947)." (R.D. Anderson, 1978)

The estimation of variance components from unbalanced data by so-called ANOVA-type methods (which are based on equating analysis of variance sums of squares of the observations to their expected values) was formally outlined by Eisenhart (1947) who also clearly distinguished between fixed- and random-effects models. Henderson (1953) considered in detail the problem of variance component estimation from unbalanced data when the underlying model was mixed. This was particularly

relevant to the estimation of variance components in animal breeding exercises, since the assumption of a mixed model is often required. A mixed model is defined and discussed in Chapter Four. Thus, the paper by Henderson (1953) outlining three methods pertinent to animal breeding is a landmark for variance component estimation (Searle, 1977).

Hartley and Rao (1967) presented the first general derivation of Maximum Likelihood (ML) for the estimation of variance components from mixed models, although Crump (1947) developed ML estimators for a specific model. Patterson and R. Thompson (1971) introduced a modified ML approach, now known as Restricted Maximum Likelihood (REML). This method follows on from the work of W.A. Thompson (1962) and also dates back to the ideas of R.L. Anderson and Bancroft (1952).

Interest in best (i.e., uniformly minimum sampling variance) quadratic, unbiased estimators (BQUE), has led to methods such as Minimum Norm Quadratic Unbiased Estimation (MINQUE) (Rao, 1970), and, under the assumption of normality, Minimum Variance Quadratic Unbiased Estimation (MIVQUE) (LaMotte, 1970).

There are certain important properties of variance component estimators that are considered to be desirable. In the discussion of Searle's (1968) paper, Zelen (1968) lists "unbiasedness, minimum variance or minimum mean square error, consistency, robustness in the sense that optimal properties do not depend critically on the distribution of the measurements, knowledge of the distribution of the estimate (either the exact or asymptotic distribution), and easy computation", as some of these features. Other favourable properties include asymptotic efficiency, minimal sufficiency and translation invariance. The prevention of estimates with values which are not

permitted by the parameter space (e.g., negative variance component estimates) is another desirable feature.

In the case of balanced data, the optimal properties of ANOVA estimators of the variance components minimizes the problem of choosing between the various methods. Graybill (1954) showed that for the general, nested, random-effects model with balanced data, that the ANOVA estimators are BQUE. Graybill and Hultquist (1961) later extended this to show ANOVA estimators are BQUE for all random-effects models with balanced data. Albert (1976) demonstrated that this property held for the mixed model case also. Furthermore, when the data are balanced and normally distributed, Graybill and Wortham (1956) showed the ANOVA estimators are best unbiased estimators (BUE).

With unbalanced data however, no one method for estimating variance components has emerged that is superior to all the other techniques in all respects (Searle, 1971b). This explains why so many methods are available for estimating variance components from unbalanced data. Moreover, the computing difficulty of many methods explains why many computing algorithms have been developed to assist in their application. Some methods for estimating variance components from unbalanced data will now be reviewed, particularly those of relevance to the present study.

#### B. Henderson's Methods

Henderson (1953) outlined three methods of estimating variance components from unbalanced data. Method I, or the ANOVA method as it is often called, was not in fact derived by Henderson but dates back to Cochran (1939). Henderson's contribution was to show that Method

I could only be applied to variance component estimation in the case of random-effects models. However, in most animal breeding applications, variance components need to be estimated from unbalanced data with an underlying mixed model. Accordingly, Henderson (1953) proposed Method II which could be applied to mixed models providing there was no nesting of random within fixed effects, or any interactions between the fixed and random effects. The limitations of Methods I and II are overcome in Method III, which yields unbiased estimates of the variance components in the general mixed model. However, the computational complexity of Method III can restrict its use.

#### 1. Henderson's Method I

With Method I, quadratic forms analogous to sums of squares in the case of balanced data are computed and then equated to their expected values. Hence the name "analysis of variance" Method. It does, however, lead to biased estimates if there are fixed effects or correlations between effects in the model. Despite this, Method I has often been used with mixed models which essentially implies either including the fixed effects in the model and treating them as random, or excluding them from the model and ignoring them completely. Searle (1968) observes that the bias is introduced because "if fixed effects are part of the model they cannot be eliminated from the expectations used in Method I."

Method I's great advantage is its computational simplicity. If a random model is justifiable, Method I is then a possible method to use. Prior to present-day computing facilities, Method I was

in fact used quite widely, but not always correctly.

## 2. Henderson's Method II

Method II was proposed to overcome the inability of Method I to handle mixed models, but to retain the computational simplicity. It involves first obtaining least squares estimates of the fixed effects in the model, adjusting the data according to these estimates and then applying Method I to the adjusted data to estimate the variance components.

Since the estimation of the fixed effects involves obtaining a generalised inverse, Searle (1968) suggested that the estimates of the variance components would depend on the particular solution vector chosen to correct the data. However, Henderson, Searle and Schaeffer (1974) subsequently showed that the variance components are, in fact, invariant to the solution vector for the fixed effects.

Method II is a possible procedure for obtaining unbiased estimates of the variance components from unbalanced data with a mixed model, notwithstanding the exceptions noted earlier.

## 3. Henderson's Method III

Method III is a general method for estimating variance components from mixed models with unbalanced data, although it is computationally more difficult to apply than Method II. Unlike the previous two methods which use ANOVA sums of squares to estimate the variance components, Method III uses reductions in sums of squares obtained from fitting the full model and different sub-models, according to the method of Yates (1934). This generates reductions in sums of squares which are free of the fixed effects

in the model. These reductions are then equated to their expected value under the full model, leading to equations which yield unbiased estimates of the variance components. As Method III is based upon the method of fitting constants by Yates (1934), it is also known as the Fitting Constants Method (Searle, 1971a). Henderson (1953) states that when it is computationally feasible, Method III is the most satisfactory of his methods for estimating variance components since it yields unbiased estimators even if certain elements in the mixed model are correlated. Its use may, however, be limited because it involves the inversion of a matrix with order equal to at least the number of random effects in the model. With animal breeding studies, for example, where hundreds or thousands of sires may be involved, this may be a problem. Despite this, Method III has been, and indeed still is, widely used particularly as computing capabilities have improved.

All three of Henderson's methods have the desirable feature of unbiasedness when applied appropriately. Another useful property is that they are translation invariant, although this is relevant to Methods II and III only. Furthermore, Henderson's methods do not require any distributional assumptions about the data.

However, Henderson's Methods are not optimal in all respects. One of the difficulties of using Henderson's Methods with non-orthogonal data is that there are often more sums of squares that can be computed than there are variance components to be estimated. Thus the choice of a different set of sums of squares will result in a different set of estimates. That is, Henderson's Methods have the undesirable property of lack of uniqueness (Henderson, 1980).

Nor do any of Henderson's Methods yield minimally sufficient statistics from unbalanced data. Furthermore, none of them possess minimum sampling variance although they all reduce to the ANOVA method (with its optimal properties) with balanced data.

The sizes of the sampling variances of the variance component estimates for a wide variety of models have been derived by many workers, but Searle (1971a) provides a comprehensive summary. He notes that these derivations often involve cumbersome algebra and the assumption of normality. In the unbalanced data case, Henderson's methods do not generate best asymptotically normal estimators.

Another disadvantage of Henderson's Methods is that they can all produce negative variance component estimates. Hudson and Van Vleck (1982) observe that "the problem of what to do when negative estimates are obtained from unbiased estimators remains unresolved." Techniques for dealing with this problem vary, although simply truncating negative estimates to zero is not a satisfactory approach (Searle, 1971b). The production of negative variance component estimates from a specific data set is often considered a criterion for using a different method of variance component estimation. So although Henderson's methods do possess the properties of unbiasedness and translation invariance, the deficiencies noted above have stimulated interest in alternative methods of variance component estimation from unbalanced data.

### C. Maximum Likelihood

The paper by Hartley and Rao (1967) led to renewed interest in using ML to estimate variance components. Harville (1977) states that "the computation of maximum likelihood estimates [of the variance

components] requires the numerical solution of a constrained non-linear optimization problem." In general terms, constrained optimization involves maximizing (or minimizing) an objective function while satisfying certain constraints. In the case of ML estimation of variance components, this specifically involves maximizing a likelihood function, subject to non-negativity constraints on the solutions obtained.

The likelihood function is defined as the joint density function,  $f_{\underline{Y}}(\underline{y}, \underline{\theta})$ , of the random variables in  $\underline{Y}$ , where the density function is known except for an unknown vector of parameters,  $\underline{\theta}$ . The method of ML yields the value of  $\underline{\theta}$  which maximizes the likelihood that  $\underline{Y}$  assumes the values of  $\underline{y}$ , the vector of observed data (Mood, Graybill and Boes, 1974). For mathematical convenience, the logarithm of the likelihood function is usually used.

When estimating variance components using ML, it is common to assume that the data are normally distributed. The necessity to make this assumption can be a limitation of ML. In animal breeding populations, the form of the distribution is rarely known, and an assumption of normally distributed residuals should not be made temerarily.

A general mixed model, defined in matrix notation is:

$$\underline{y} = \underline{X}\underline{\alpha} + \sum_{i=0}^c \underline{Z}_i \underline{u}_i \quad (2.1)$$

where,

$\underline{y}$  is an  $N \times 1$  vector of observations

$\underline{X}$  is a known  $N \times p$  design matrix

$\underline{\alpha}$  is an unknown  $p \times 1$  vector of fixed effects

$\underline{Z}_i$  for  $i = 0, 1, \dots, c$  are known  $N \times q_i$  incidence matrices

$\underline{u}_i$  for  $i = 0, 1, \dots, c$  are non-observable  $q_i \times 1$  vectors of random effects

such that,

$$E [u_i] = 0, i = 0, 1, \dots, c$$

with the elements of  $u_i$  assumed to be independently, normally distributed with common variance  $\sigma_i^2$ ,  $i = 0, 1, \dots, c$ . It is also assumed that  $\text{cov}(u_i, u_{i'}) = 0$ ,  $i \neq i'$  and therefore,

$$E [y] = X\alpha \text{ and } \text{var}(y) = V = \sum_{i=1}^c \sigma_i^2 Z_i Z_i'$$

where  $y$  is assumed to have a multivariate normal distribution.

For the model (2.1), the likelihood function can then be derived under normality (see Hartley and Rao, 1967) as:

$$L = \frac{1}{2} \log 2\pi - \frac{1}{2} \log [V] - \frac{1}{2} (y - X\hat{\alpha})' V^{-1} (y - X\hat{\alpha}) \quad (2.2)$$

Deriving equations for "estimating"  $\alpha$ , the vector of fixed effects and  $\hat{\sigma}^2$ , the vector of variance components pertaining to the model, where,

$$\hat{\sigma}^2 = [\sigma_0^2 \sigma^2] \text{ with } \sigma^2 = \{\sigma_i^2\}, i = 1, \dots, c \quad (2.3)$$

involves differentiating (2.2) with respect of  $\alpha$  and  $\hat{\sigma}^2$  and equating to zero. This leads to the equations:

$$\tilde{\alpha} = (X' \tilde{V}^{-1} X)^{-1} X' \tilde{V}^{-1} y \quad (2.4)$$

$$(\text{tr} [\tilde{V}^{-1} Z_i Z_i' \tilde{V}^{-1} Z_j Z_j']) \tilde{\sigma}^2 = (y' \tilde{P} Z_i Z_i' \tilde{P} y) \quad (2.5)$$

for  $i, j = 0, 1, \dots, c$

with  $\tilde{\alpha}$  being a solution for  $\alpha$  and where  $\tilde{P}$  is  $P$  of (2.6), but with  $\tilde{V}$  in place of  $V$ :

$$\tilde{P} = \tilde{V}^{-1} - \tilde{V}^{-1} X (X' \tilde{V}^{-1} X)^{-1} X' \tilde{V}^{-1} \quad (2.6)$$

where  $\tilde{V}$  is a ML solution for  $V$ .

Note that these ML equations require  $V$  to be known which is unlikely. Nevertheless, assuming an initial value for  $\tilde{\sigma}^2$ , a solution for  $\hat{\sigma}^2$  may be obtained iteratively. Note that only non-negative solutions to (2.4) are ML estimators of  $\hat{\sigma}^2$ .

The ML equations do not usually yield explicit solutions. Exceptions include certain balanced cross-classified models with as many latent roots as there are variance components to be estimated (T.W. Anderson, 1969) and all completely nested models with balanced data (Szatrowski and Miller, 1977). Iterative algorithms do not always lead to convergence and may be tedious. However with recent improved computing facilities and the derivation of improved computing algorithms, ML estimation has become less difficult to apply. Henderson (1980) observes that the iteration can still be slow, and that even when convergence does occur fairly rapidly that there is no guarantee that this is to a global rather than a local maximum. This latter aspect however does depend on the computing algorithm used, which is further considered in section F. It should be noted that applying procedures (such as penalty techniques or gradient projection methods) to deal with convergence problems will involve a large, and often impractical, increase in computing.

Hartley and Rao (1967) also addressed some of the optimal properties of ML estimation. They used the Methods of Steepest Ascent (see section F) for the iteration, which guarantees convergence and non-negative estimates, although it is often impractically slow. Hartley and Rao concentrated primarily on large sample optimality criteria. They showed that for any mixed model with unbalanced data, that the ML estimator they derived is weakly consistent and best, asymptotically normal. Hartley and Rao assumed the "design" matrices  $\underline{X}$  and  $\underline{Z}$  had full rank and that the number of levels of each random factor is restricted to some constant, however T.W. Anderson (1973) and Miller (1973, see also Miller, 1977) showed that ML estimators were consistent

and best asymptotically normal without the second of these assumptions. Hartley and Rao also noted that, based on work by Hultquist and Graybill (1965) that "it is a considerable small sample advantage" that ML estimators will always be functions of minimally sufficient statistics.

Searle (1970, 1971a) derived the large sample variances of ML estimators for a number of models. Since ML estimators are best asymptotically normal he notes that "one value of the resulting expressions is that even though ML estimators of the  $\sigma^2$ 's cannot be obtained [explicitly], values for the variances in their asymptotic distribution can; and against these can be compared variances of estimators obtained by other methods to give measures of asymptotic efficiency of those other methods."

The main difficulty with the ML techniques proposed by Hartley and Rao (1967) is the amount of computation time. Hence they also discussed an alternative method which was further developed by Henderson (1973) and Henderson, Ufford and Schaeffer (1976). This ML method is known as Henderson's Algorithm and makes use of Henderson's Mixed Model Equations (MME) (see, for example, Henderson *et al.*, 1959; Henderson, 1963, 1972). This greatly simplifies the computations. Henderson's Algorithm also has the advantage that it always yields non-negative variance component estimates. Yet another algorithm was developed by T.W. Anderson (1973), although it does not guarantee non-negative solutions. There are also a number of general algorithms which have been applied and some of these, e.g. Fisher-scoring and Newton-Raphson, are discussed in section F.

Harville (1977) provides a comprehensive review of ML methodology

and properties. He notes that the ML approach is always well-defined as well as having the optimal properties of consistency, sufficiency and efficiency already discussed.

An important problem associated with ML when estimating variance components from mixed models is that no account is taken of the loss in degrees of freedom due to the need to also estimate the fixed effects, in the model. Henderson (1980) states that unless the rank of the design matrix,  $X$ , pertaining to the fixed effects is small relative to the total number of observations, the estimate of the residual variance component will be seriously biased downwards. He also notes that the estimate of the variance component pertaining to any random effect with a small number of levels will also be biased downwards. To overcome this problem, Patterson and Thompson (1971) developed the REML method discussed in the following section D.

Harville (1977) suggests that another disadvantage of ML, namely the need to assume a distributional form, is not in fact very serious. He discusses several results which indicate that ML estimators derived under normality may be suitable even when the form of the distribution is not specified. Thus, except where the estimators are biased by fixed effects in a mixed model, the advantages of ML have recently seen it applied in many appropriate situations. For example, Rothschild and Henderson (1979) and Rothschild, Henderson and Quaas (1979) have used ML to estimate variance and covariance components from simulated lactation records; Schaeffer (1979, 1981) applied it to the estimation of variance and covariance components from pig data and dairy sire progeny test results, respectively.

D. Restricted Maximum Likelihood

Restricted Maximum Likelihood (REML), being a modification of ML, shares many of the properties of ML. The main advantage of REML for variance component estimation is that it takes account of the need to estimate fixed effects in mixed models. The method is described by Patterson and Thompson (1971) essentially as partitioning the data into two separate parts and deriving the likelihood function for each. One portion, which is free of the fixed effects is then maximized to estimate the vector of variance components  $\hat{\sigma}^2$ , defined in (2.3).

Quaas (1976) suggests that the development of the REML equations can be considered in two different ways. In the first approach, the equations are derived by maximizing the likelihood function over a restricted sample space. That is the likelihood is derived for just a portion of the sample space of  $\underline{y}$ , the observation vector. To obtain variance component estimates by REML, the likelihood function is derived from "a certain subpopulation for which  $[\hat{\alpha}]$  has a fixed value" (Quaas, 1976). Thus, the REML likelihood function is the conditional density of  $\underline{y}$  given  $\hat{\alpha}$  which is assumed to have a multivariate normal distribution with  $E[\underline{y}|\hat{\alpha}] = \hat{\alpha}$  and  $\text{var}(\hat{y}|\hat{\alpha}) = \underline{V} - \underline{X}(\underline{X}'\underline{V}^{-1}\underline{X})^{-1}\underline{X}'$ .

The second approach is the method of maximizing the likelihood of a set of error contrasts. An error contrast is a linear function of the observation vector, say  $\underline{k}'\underline{y}$ , where  $\underline{k}$  does not include any of the unknown parameters in the model and  $E[\underline{k}'\underline{y}] = 0$ . This is the method used by Patterson and Thompson (1971) to derive their REML equations. The maximum number of linearly independent (LIN) contrasts in each set is  $N - p^*$  where  $N$  is the total number of observations (i.e., the order of  $\underline{y}$ ) and  $p^*$  is the rank of  $\underline{X}$ , the design matrix. Searle (1979a)

states that "there is obviously no merit in dealing with  $\underline{k}'\underline{y}$  if some rows of  $\underline{k}'$  (and hence elements of  $\underline{k}'\underline{y}$ ) are linear combinations of others; neither should we lose information by using a  $\underline{k}'\underline{y}$  that has fewer elements than the possible maximum."  $\underline{k}'\underline{y}$  has a multivariate normal distribution with  $E[\underline{k}'\underline{y}] = 0$  and  $\text{var}(\underline{k}'\underline{y}) = \underline{k}'\underline{V}\underline{k}$ . The logarithm of the likelihood of  $\underline{k}'\underline{y}$  is then required, and it can be shown (see Searle, 1979b, for example) for any  $\underline{k}'$  of order  $(N - p^*) \times N$ , with full row rank  $(N - p^*)$ , and with  $\underline{k}'\underline{X} = 0$ , that the resulting function will be exactly the same except for a constant which depends only on  $\underline{k}'$  and not on any of the unknown parameters in the model. Quaas (1976) has shown that maximization over a restricted sample space yields exactly this expression for the likelihood. Thus the likelihood function for REML is:

$$L_1 = \text{constant} - \frac{1}{2} \log[\underline{V}] - \frac{1}{2} \log[\underline{X}'\underline{V}^{-1}\underline{X}] - \frac{1}{2} (\underline{y} - \underline{X}\hat{\underline{\alpha}})' \underline{V}^{-1} (\underline{y} - \underline{X}\hat{\underline{\alpha}}) \quad (2.7)$$

Various procedures can be used to obtain the REML equations from  $L_1$ . They generally involve various combinations of taking first and second order partial derivatives of (2.7), their expected values and/or related quantities (Harville, 1977). One way to obtain the REML equations is to maximize the likelihood by differentiating  $L_1$  with respect to  $\underline{\theta}$  and then setting the resulting equations to zero, in the same way that the equations for ML were obtained. From Searle (1979a), the equations which afford REML estimates are:

$$\text{tr}(\hat{\underline{P}}\underline{Z}_i\underline{Z}_i') = \underline{y}'\hat{\underline{P}}\underline{Z}_i\underline{Z}_i'\hat{\underline{P}}\underline{y} \quad (2.8)$$

$$\text{tr}(\hat{\underline{P}}\underline{Z}_i\underline{Z}_i') = (\underline{y} - \underline{X}\hat{\underline{\alpha}})' \hat{\underline{V}}^{-1} \underline{Z}_i\underline{Z}_i' \hat{\underline{V}}^{-1} (\underline{y} - \underline{X}\hat{\underline{\alpha}}) \quad (2.9)$$

or,

$$(\text{tr}[\hat{\underline{P}}\underline{Z}_i\underline{Z}_i'\hat{\underline{P}}\underline{Z}_j\underline{Z}_j']) \hat{\underline{\sigma}}^2 = (\underline{y}'\hat{\underline{P}}\underline{Z}_i\underline{Z}_i'\hat{\underline{P}}\underline{y}) \quad (2.10)$$

for  $i, j = 0, 1, \dots, c$

where the matrix  $\hat{P}$  is the matrix  $P$  of (2.6) with  $\hat{V}^{-1}$  in place of  $V^{-1}$ ,  $\hat{\alpha}$  is a solution for  $\alpha$  and  $\hat{\sigma}^2$  is the REML estimate of  $\sigma^2$ .

As Searle (1979b) observes, while there is no implicit REML estimator of  $\alpha$  obtained from maximizing (2.7) an "obvious" estimator of  $\alpha$  is:

$$\hat{\alpha} = (\tilde{X}'\hat{V}^{-1}\tilde{X})^{-1}\tilde{X}'\hat{V}^{-1}y \quad (2.11)$$

where  $\hat{\alpha}$  is, in fact, the generalised least squares (GLS) estimate  $\hat{\alpha} = (\tilde{X}'\tilde{V}^{-1}\tilde{X})^{-1}\tilde{X}'\tilde{V}^{-1}y$ , but with  $\hat{V}^{-1}$  in place of  $\tilde{V}^{-1}$ . Note also, that  $\hat{\alpha}$  is analogous to  $\tilde{\alpha}$  of (2.4) with  $\tilde{V}^{-1}$  replaced by  $\hat{V}^{-1}$ .

Maximization over a restricted sample space is therefore the basic feature distinguishing REML from ML. The REML equations do not guarantee non-negative solutions, so like ML, only non-negative solutions to (2.8) - (2.10) can be regarded as REML estimates of the variance components. As with ML, the REML equations do not usually yield explicit estimators but need to be solved iteratively. Various computing algorithms are available which are discussed in more detail in the following section E. Some use second order partial derivatives or expected values, and they differ in speed of convergence and other properties. For example, an algorithm may be modified to avoid negative estimates.

Another feature of REML is that with balanced data, solutions to the REML equations have been shown to be numerically equal to ANOVA estimates in special cases including the split-plot design (Patterson and Thompson, 1974), the 1-way random, the 2-way nested random and the 2-way crossed classification mixed model, with and without interaction (Corbeil and Searle, 1976b). R.D. Anderson (1978) in fact showed that in the balanced data case, REML estimators of the

variance components coincide with ANOVA estimators in general. That is, with balanced data REML estimators are BUE.

REML does not improve on ML in all properties of interest. Although it should be pointed out that, to date, the properties of ML estimators have been studied far more extensively than those of REML estimators. One property in which ML is superior is sufficiency. In the comment to Harville's (1977) paper, Rao (1977) observes that in contrast to the full ML likelihood function, the REML likelihood function is not minimally sufficient. In his rejoinder Harville suggests that this is irrelevant since "the ML estimator turns out to depend on the data vector  $y$  only through the error contrasts."

REML estimators of the variance components are translation invariant (Corbeil and Searle, 1976a). Harville (1977) notes that this has been cited as a reason for preferring REML over ML, but that this is quite invalid since ML estimators are themselves translation invariant also.

The relative sampling variance of the REML estimator depends largely on the "specifics of the underlying model and possibly on  $\underline{\theta}^+$ " (Harville, 1977), where  $\underline{\theta}^+$  is the true value of  $\underline{\theta}$ . The sampling variance of REML in comparison to ML estimators has been studied for ordinary fixed ANOVA or regression models. For such models Harville (1977) states that "the ML estimator of the variance  $[\underline{\theta}_0]$  has uniformly smaller MSE than the REML estimator when  $p^* \leq 4$ ; however, the REML estimator has the smaller MSE when  $p^* \geq 5$  and  $n - p^*$  is sufficiently large ( $n - p^* > 2$  suffices if  $p^* \geq 13$ ). MSE comparisons between variance-component estimators obtained by solving the likelihood equations for ML and variance-component estimators obtained by solving the likelihood

equations for REML, were made by Corbeil and Searle (1976b) and Hocking and Kutner (1975) for several mixed and random ANOVA models." They also found that the ML estimator generally had a smaller MSE relative to the REML estimator except when there were only a few observations per cell or when the random-effects variances are small relative to the residual variance, which Corbeil and Searle observe is often the case in genetics.

Corbeil and Searle (1976b) also investigated the relative properties of ML, REML, Henderson's Method II and Method III (Henderson, 1953) and an iterative method from Thompson (1969) in the unbalanced data case using numerical studies. They found ML to be noticeably superior in terms of MSE, the other methods being relatively similar. The same trend was found for the sampling variances which were calculated from expressions derived by Hartley and Rao (1967) and Corbeil and Searle (1976a). However, Corbeil and Searle (1976b) do note that the study was limited to a mixed model with 6 levels of the fixed effects and 10 levels of the random effects factor and that these results may not necessarily apply with larger models.

Carter (1981) looked at some large sample properties of REML but only for the regression model. He was interested in comparing different methods of measuring the activity of the enzyme sucrose in intestinal tissues, using REML method. Carter first studied some asymptotic properties of REML estimators under a full rank model. He found them to be consistent and asymptotically, normally efficient when just 3 levels (i.e. 3 different measuring methods) of the regression coefficients are included in the model. Nevertheless, as both Harville and Rao (in Harville, 1977) point out, there is still a great need for

"realistic" large-sample properties of REML estimators to be established for a wide variety of models. Indeed, work on asymptotic properties of ML estimators also needs to be extended to a wider range of models than have currently been investigated.

The usefulness of any method of estimation depends finally on the ease with which it can be applied. Henderson (1980) comments that the iteration is usually more complex with REML than it is for ML. Nevertheless, there are a number of algorithms that can be applied to REML to improve convergence, which will be discussed in the subsequent section. The choice of an appropriate algorithm and the availability of suitable computing facilities makes REML an applicable method of estimation in a variety of situations.

#### E. Other Estimation Methods

Attention to the properties of unbiasedness and minimum sampling variance has stimulated interest in best, quadratic unbiased estimation (BQUE). Townsend (1968) and Townsend and Searle (1971) derived estimators for a 1-way random model with unbalanced data, assuming a normal distribution with a zero mean, that were locally BQUE, where locally best is defined by LaMotte (1973a). Harville (1969) also derived locally BQUE of the variance components under normality for a 1-way random model and unbalanced data, but for the case of non-zero means.

The BQUE approach has been considerably extended by the work of LaMotte (1970, 1973a, b) and Rao (1970, 1971a, b, 1972, 1973). The minimum norm quadratic unbiased estimators (MINQUE) of the variance components were derived by Rao (1970) by minimizing a

weighted, Euclidian norm (as discussed, for example, by R.D. Anderson, 1978, 1979b and Searle, 1979b). The MINQUE do not require an iterative solution and are distribution free.

Under the assumption of normality, LaMotte (1970) derived minimum variance quadratic unbiased estimators (MIVQUE) which are locally BQUE. The MIVQUE also do not require an iterative solution. In fact, under normality the form of the MINQUE equations is identical to the MIVQUE equations for estimating the variance components:

$$(\text{tr}[\hat{P}Z_i Z_i' \hat{P}Z_j Z_j']) \hat{\sigma}^2 = (y' \hat{P}Z_i Z_i' \hat{P}y) \quad (2.12)$$

for  $i, j = 0, 1, \dots, c$

where  $\hat{P}$  is  $P$  of (2.6) but with a priori estimates of the variance components,  $\hat{\sigma}^2$ , in  $V$ . Equations (2.12) also have the same form as equations (2.8) - (2.10) for REML, although REML requires an iterative solution.

The properties of translation invariance and unbiasedness of the estimator are invoked for MINQUE and MIVQUE. Denoting a quadratic estimator of the estimable function,  $t' \hat{\sigma}^2$ , as  $y' \underline{A} y$  for a symmetric matrix  $\underline{A}$ , the necessary and sufficient conditions for translation invariance are that  $\underline{A} \underline{X} = 0$ . Unbiasedness is guaranteed when  $\underline{X}' \underline{A} \underline{X} = 0$  and  $\text{tr}(\underline{A} Z_i Z_i') = t_i$ ,  $i = 0, 1, \dots, c$ , which ensures that  $E[y' \underline{A} y] = t' \hat{\sigma}^2$ .

In the derivation of MIVQUE, minimum sampling variance of the estimator is also invoked. That is,  $\underline{A}$  must also be determined such that the variance of  $t' \hat{\sigma}^2$  is minimized. R.D. Anderson (1978) notes that the estimator  $t' \hat{\sigma}^2$  is only MIVQUE when the chosen a priori vector of the variance components  $\hat{\sigma}^2$ , actually equals  $\sigma^2$ , the variance components to be estimated. Since the probability that the a priori estimates do equal  $\hat{\sigma}^2$  is low, Anderson queries whether the minimum variance properties of MIVQUE or MINQUE, under normality, are likely to

be very strong in practice for a mixed model and unbalanced data.

R.D. Anderson, H.V. Henderson, Pukelshein and Searle (1979) have shown that, with no assumptions concerning the distribution, the MINQUE are BQUE with balanced data.

There are a number of variations of MINQUE and MIVQUE which are less dependent on a priori estimates of  $\dot{\sigma}^2$ . Rao (1970) suggested that one method of overcoming lack of knowledge about  $\dot{\sigma}^2$ , is to set all the variances in  $\dot{\sigma}^2$  to zero and assume any value for  $\sigma_0^2$ . This method is now known as MINQUE-(0) and yields the same estimators as were derived by Seely (1971) and Hartley, Rao and LaMotte (1978). Assuming normality, Quaas and Bolgiano (1979) showed that the sampling variance of the MIVQUE-(0) estimators became very large as  $\dot{\sigma}^2/\sigma_0^2$  deviated from 0.

Another modification of MINQUE which alleviates the need for a priori estimates of the variance components is iterative or I-MINQUE, so-named by Brown (1976). He also showed that both the MINQUE and I-MINQUE estimators are asymptotically normal. I-MINQUE is computed by iterating the MINQUE equations (2.12) starting with any reasonable values for  $\dot{\sigma}^2$ . However, the estimators are no longer unbiased nor do they possess minimum variance. Thus, Searle (1979b) suggests that the loss of minimum variance properties makes an analogous procedure for MIVQUE inappropriate.

Under the assumption of normality, the first iterate of the REML or the I-MINQUE procedures equals the MINQUE and the MIVQUE (R.D. Anderson, 1978). As with the REML solutions, all the MINQUE-related methods discussed must be constrained to non-negative solutions when estimating variance components P.S.R.S. Rao (1977) discusses a

minimum norm quadratic estimator (MINQE) which is inherently non-negative. However, it is not unbiased, although its sampling variance was found to be comparable to the MINQE.

There are various other modifications of MINQE available. Rao (1972) extended the use of MINQE for variance components to include the estimation of covariance components. Goodnight (1976) employed MIVQUE "to provide quadratic unbiased estimators whose variance is dependent on as few of the unknown variance components as possible", which he called maximally-invariant quadratic unbiased estimators. Harville (1981) derives the necessary and sufficient conditions for unbiased and minimum-variance unbiased estimators for fixed effect models with arbitrary covariance structure. These estimators have also been studied by Rao (1979).

Henderson (1980) describes a new method of estimation of variance components applicable to mixed models which is related to MINQE (0) but does not depend on  $\sigma/\sigma_0^2$  being close to 0. Furthermore, it is computationally feasible, with  $\underline{X}'\underline{X}$  being the only matrix to be inverted, although this is also unnecessary if  $\underline{X}'\underline{X}$  is diagonal. Henderson notes that the sampling variances are relatively easy to compute for this method. He thus compared these with MINQE, MINQE (0) and Method III for a mixed model including one random (sire) effect without interaction, over a wide range of  $\sigma_0^2/\sigma_s^2$  values. The new method showed a smaller sampling variance than all the other methods over the entire range of  $\sigma_0^2/\sigma_s^2$  studied (from 0.5 to 100). Hudson and Van Vleck (1982) compared Henderson's new method with Method III. The estimates resulting from the two methods were generally similar although they occasionally yielded estimates of opposite sign. The possible production of

negative estimates is an unresolved problem of both methods. Both methods are unbiased and translation invariant. The ease of computation of the new methods is improved when there are fewer fixed effect equations to be absorbed although both methods are generally computationally feasible with present-day facilities. Once the fixed effects are absorbed the new method converges rapidly and with relative ease. Thus, Henderson's new method appears to have a number of desirable properties for estimating variance components in animal breeding settings. Nevertheless, a universally optimal method is not available for unbalanced data and there are methods other than those described which are desirable for particular applications.

#### F. Computing Algorithms

Having derived equations for a particular method of estimation it is then necessary to obtain a solution. Attention will be focussed here on computing algorithms that can be applied to REML. The computing algorithm may involve solving the equations of (2.8) - (2.10), or it could incorporate modifications which offer advantages such as speeding up the iteration or yielding only solutions which are permitted by the parameter space.

##### 1. Mathematical optimization

Before discussing some of the algorithms, a brief description of the terminology of mathematical optimization will be given. Mathematical optimization algorithms can be single-variate or multivariate, constrained or unconstrained, direct search methods or indirect gradient methods. Gradient methods are based on the Taylor series and tend to be more useful for variance component

estimation because search methods are usually very time-consuming.

Harville (1977) states that the solution for  $\underline{\theta}$ , the vector of unknown parameters from the  $(r + 1)$ st iteration of any maximization algorithm can be represented as:

$$\tilde{\underline{\theta}}^{(r+1)} = \tilde{\underline{\theta}}^r + p_r \underline{w}_r \quad (2.13)$$

where  $\tilde{\underline{\theta}}^r$  is the solution for  $\underline{\theta}$  from the previous round,  $\underline{w}_r$  is the vector determining the search direction, and  $p_r$  is a positive scalar controlling the distance traversed in each round, in the direction of the search. Gradient algorithms are further defined, for REML, by specifying for  $L_1$ , the likelihood function under REML, that:

$$\underline{w}_r = \underline{N}_r (\partial L_1 / \partial \underline{\theta})^{(r)} \quad (2.14)$$

for some square matrix  $\underline{N}_r$  with order equal to the number of random effects in the model. Equation (2.14) could just as easily have been written for ML, that is, with  $\partial L / \partial \underline{\theta}$ .

The different gradient algorithms are characterized by the terms they have in  $p_r$  and  $\underline{N}_r$ . A method may be either first order or second order depending on whether or not it uses the second order partial derivatives of the likelihood function as well as the first order partial derivatives in  $\underline{N}_r$ . The steepest ascent method (used by Hartley and Rao, 1967) is a first order method, while the Fisher-scoring and Newton-Raphson algorithms are both second order methods.

The constrained optimization methods usually apply either linear or non-linear constraints, although with some algorithms (e.g., the gradient projection and the penalty techniques), both can be applied. For reasons of mathematical simplicity, it is generally preferable

to apply linear constraints. There are also transformation methods available which change a constrained problem into an unconstrained one. Note, that "non-linear optimization method" refers to a non-linear objective function (e.g., quadratic for variance components); it does not specify the type of constraints. It is also important to recognize that for any given estimation method, including REML, it is not possible to identify the optimal computing algorithm. An algorithm that performs well in one application may perform poorly in another setting. The rate at which the iteration converges and certain optimality considerations (such as whether the iteration converges to a global rather than a local maximum, or whether only solutions permitted by the parameter space can be generated), have resulted in the derivation of a variety of algorithms. Thus it is useful to have some knowledge of the performance that can be expected from various algorithms under specific conditions.

## 2. Unconstrained optimization algorithms

The method of steepest ascent (also known as Cauchy's method) used by Hartley and Rao (1967) is a first order method. The iteration may be initiated with any set of starting values and it is theoretically guaranteed to converge. However, it has the disadvantage of having a linear convergence rate and is often very slow to converge (Powell, 1970). This tedious rate is associated with the matrix of order  $N$ , the number of observations, that has to be inverted on each round of the iteration. Hartley and Rao also outline a modification which Hemmerle and Hartley (1973) developed into an algorithm called the  $W$  transformation, which can

be applied to other algorithms such as the steepest ascent method. It removes the need to invert this  $N \times N$  matrix and ensures the computations do not depend on  $N$  at all, thus speeding up the iteration. Hemmerle and Hartley actually apply the  $W$  transformation to the Newton-Raphson algorithm. This is one of two important gradient algorithms, the other being Fisher-scoring, which will be discussed here.

The method of Fisher-scoring is a second-order gradient method. It also has a linear convergence rate, however, it is faster than the steepest ascent method since it does not involve inverting an  $N \times N$  matrix (where  $N$  is the total number of observations), on each iteration. With this method,  $p_r$  of (2.13) equals 1 and  $\tilde{N}_r$  of (2.14) is a function of the expected values of the second order partial derivatives. Thus, based on Searle (1979b) the solution for  $\underline{\theta}$  from the  $(r + 1)$ st iteration of applying the Fisher-scoring algorithm for REML is:

$$\underline{\theta}^{(r+1)} = \tilde{\underline{\theta}}^r + - E \left[ \frac{\partial^2 L_1}{\partial \underline{\theta} \partial \underline{\theta}'} \right]^{-1} \frac{\partial L_1}{\partial \underline{\theta}} \quad (2.15)$$

To express (2.15) in a computational form, the information matrix of  $\underline{\theta}$ , denoted  $I(\underline{\theta})$ , is first defined (from Searle, 1979b), under REML as:

$$I(\underline{\theta}) = - E \left[ \frac{\partial^2 L_1}{\partial \underline{\theta} \partial \underline{\theta}'} \right] \quad (2.16)$$

Thus  $I(\underline{\theta})$  is used directly in the Fisher-scoring algorithm. It is often useful to denote  $\underline{\theta}$  as a function of  $\dot{\sigma}^2$ , defined in (2.2), say  $\dot{\underline{\gamma}}$ , where:

$$\dot{\underline{\gamma}} = [\underline{\gamma}_0 \underline{\gamma}] \quad (2.17)$$

for  $\gamma_0 = \sigma_0^2$  and  $\underline{\gamma} = \{\gamma_i\} = \{\sigma_i^2/\sigma_0^2\}$  for  $i, = 1, \dots, c$ . Then the expression for  $I(\dot{\underline{\gamma}})$  may be given as:

$$I(\dot{\underline{\gamma}}) = \frac{1}{2} \begin{bmatrix} \{(N-p^*)/\sigma_0^4\} & \{((q_i - \text{tr}[T_{ii}])/\sigma_i^2)'\} \\ & \text{diagonal terms:} \\ & \text{tr}[(I_{q_i} - T_{ii})^2]/\gamma_i^2 \\ \{(q_i - \text{tr}[T_{ii}])/\sigma_i^2\} & \text{off-diagonal terms:} \\ & \text{tr}[T_{ij}T_{ji}]/\gamma_i\gamma_j \end{bmatrix} \quad (2.18)$$

for  $i, j = 1, \dots, c$

which may be substituted into (2.15) and the remaining expression for the first order partial derivatives for  $\dot{\underline{\gamma}}$  from Harville (1977) is computed by the Fisher-scoring algorithm as:

$$\frac{\partial L_1}{\partial \dot{\underline{\gamma}}} = -\frac{1}{2} \begin{bmatrix} \{\gamma_0^{-1}(N-p^* - \underline{y}'P'\underline{y})\} \\ \{\gamma_i^{-1}(q_i - \text{tr}(T_{ii})) - \gamma_0 v_i' v_i\} \end{bmatrix} \quad (2.19)$$

for  $i = 1, \dots, c$

where  $P$  is defined in (2.6),  $q_i$  is the number of levels in the  $i$ th random effect, and  $v_i$  is an element of  $\underline{v}$  which is related to  $\underline{b}$ , the vector of random effects:

$$\underline{b} = D\underline{v} \quad (2.20)$$

where  $D = \text{var}(\underline{b})$ .  $T_{ii}$  is a submatrix of  $T$ :

$$\underline{T} = (\underline{I} + \underline{Z}'\underline{S}\underline{Z})^{-1} = \begin{bmatrix} T_{11} & \cdot & \cdot & \cdot & T_{1c} \\ \vdots & \cdot & & & \vdots \\ T_{c1} & \cdot & \cdot & \cdot & T_{cc} \end{bmatrix} \quad (2.21)$$

with

$$\underline{S} = \underline{R}^{-1} - \underline{R}^{-1}\underline{X}(\underline{X}'\underline{R}^{-1}\underline{X})^{-1}\underline{X}'\underline{R}^{-1} \quad (2.22)$$

solving (2.19) therefore requires estimates of  $\underline{P}$ ,  $\underline{T}$  and  $\underline{v}$ .

Solutions for  $\underline{T}$  and  $\underline{v}$  are discussed in the following section, F3.

Searle (1979b) has shown that:

$$\underline{y}'\underline{P}\underline{y} = \underline{y}'\underline{S}(\underline{y} - \underline{Z}\underline{D}\underline{v}) \quad (2.23)$$

This expression can be further simplified depending on the particular specifications of the underlying model. For example, when  $\underline{R} = \begin{bmatrix} \underline{I} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix}$  can be assumed then:

$$\underline{y}'\underline{P}\underline{y} = [(\underline{y}'\underline{y} - \underline{y}'\underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{y})/\sigma_0^2] - [(\underline{y}'\underline{Z} - \underline{y}'\underline{X}(\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Z})\underline{v}\underline{v}'] \quad (2.24)$$

The Newton-Raphson algorithm is, in fact, very similar to the Fisher-scoring method except that the second order partial derivatives themselves are used, rather than their expected values. Thus, it is also a second order gradient method. From Searle (1979b), the REML solution to  $\underline{\Theta}$  on the  $(r+1)$ st iteration of the Newton-Raphson algorithm is:

$$\underline{\Theta}^{(r+1)} = \underline{\Theta}^{(r)} - \left[ \frac{\partial^2 L_1}{\partial \underline{\Theta} \partial \underline{\Theta}'} \right]^{-1} \left[ \frac{\partial L_1}{\partial \underline{\Theta}} \right] \quad (2.25)$$

Computing the first order term in (2.25) is exactly the same as the expression given in (2.19) for the Fisher-scoring method. The computation of the second order term is given by:

$$\frac{\partial^2 L_1}{\partial \underline{\gamma} \partial \underline{\gamma}'} = -\frac{1}{2} \left[ \begin{array}{l} \{\underline{\gamma}_0^{-2}(-N+p^*+2\underline{y}'\underline{P}\underline{y})\} \quad \{(\underline{\tilde{v}}_i' \underline{\tilde{v}}_i)'\} \\ \text{diagonal terms:} \\ -\underline{\gamma}_i^{-2} \text{tr}(\underline{I}_{qi} - \underline{T}_{ii})^2 + 2\underline{\gamma}_0 \underline{\gamma}_i^{-1} \underline{\tilde{v}}_i' (\underline{I}_{qi} - \underline{T}_{ii}) \underline{\tilde{v}}_i \\ \text{off-diagonal terms:} \\ -\underline{\gamma}_i^{-1} \underline{\gamma}_j^{-1} \text{tr}(\underline{T}_{ij} \underline{T}_{ji}) - 2\underline{\gamma}_0 \underline{\gamma}_j^{-1} \underline{\tilde{v}}_i' \underline{T}_{ij} \underline{\tilde{v}}_j \end{array} \right] \quad (2.26)$$

for  $i, j = 1, \dots, c$

where  $\underline{y}'\underline{P}\underline{y}$  is defined as in (2.23) or, for  $\underline{R} = \underline{I}\sigma_0^2$ , as in (2.24).

The advantage of the Newton-Raphson algorithm is that it has a quadratic convergence rate and therefore a solution is usually obtained in fewer iterations than from the Fisher-scoring method. Harville (1977) notes that for any function with a negative definite Hessian matrix (the matrix of second order partial derivatives of the objective function), convergence will occur in a single iteration. The difficulty with the Newton-Raphson algorithm is that quadratic convergence rates are only achieved if the iteration is initiated with starting values that are in reasonable proximity to the maximizing value. If the starting value is poor, the Newton-Raphson algorithm may converge to a local rather than a global maximum, or to a stationary point which is neither, or it may fail to converge altogether (Bard, 1974).

When estimating variance components, well-substantiated prior estimates are usually available as suitable starting values for the Newton-Raphson algorithm. Jennrich and Sampson (1976) found the performance of the Newton-Raphson iteration improved if it was preceded by at least one round of Fisher-scoring and also defaulted to this whenever the change in  $\hat{\gamma}_i$  was greater than 1. Harville (1977) suggests that although the Fisher-scoring method usually takes more iterations to converge, that it is often easier, and quicker, to compute the terms required within each iteration. Thus both methods have computational attributes.

### 3. Use of the mixed model equations

When REML estimates of the variance components are to be generated from a mixed model with unbalanced data, it is generally an advantage to make use of the mixed model equations (MME). The

MME derived by Henderson (1963, for example) are:

$$\begin{bmatrix} \tilde{X}'\tilde{R}^{-1}\tilde{X} & \tilde{X}'\tilde{R}^{-1}\tilde{X} \\ \tilde{Z}'\tilde{R}^{-1}\tilde{X} & \tilde{Z}'\tilde{R}^{-1}\tilde{Z} + \tilde{D}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\tilde{\alpha}} \\ \tilde{b} \end{bmatrix} = \begin{bmatrix} \tilde{X}'\tilde{R}^{-1}\tilde{y} \\ \tilde{Z}'\tilde{R}^{-1}\tilde{y} \end{bmatrix} \quad (2.27)$$

Harville (1976, 1977) suggested another form of the MME which, unlike Henderson's, allowed for singular,  $\tilde{D}$ , namely:

$$\begin{bmatrix} \tilde{X}'\tilde{R}^{-1}\tilde{X} & \tilde{X}'\tilde{R}^{-1}\tilde{Z}\tilde{D} \\ \tilde{Z}'\tilde{R}^{-1}\tilde{X} & \tilde{I} + \tilde{Z}'\tilde{R}^{-1}\tilde{Z}\tilde{D} \end{bmatrix} \begin{bmatrix} \hat{\tilde{\alpha}} \\ \tilde{v} \end{bmatrix} = \begin{bmatrix} \tilde{X}'\tilde{R}^{-1}\tilde{y} \\ \tilde{Z}'\tilde{R}^{-1}\tilde{y} \end{bmatrix} \quad (2.28)$$

Obtaining a solution for  $\tilde{v}$  from (2.28) and making use of the definition of  $\tilde{v}$  in (2.20), provides a solution for  $\tilde{b}$ . The left hand side (LHS) matrix of (2.28), say  $\tilde{C}$ , is related to the LHS matrix of (2.27), say  $\tilde{B}$ , by the expression:

$$\tilde{C} = \tilde{B} \begin{bmatrix} \tilde{I} & \tilde{0} \\ \tilde{0} & \tilde{D} \end{bmatrix} \quad (2.29)$$

In comparison with the matrices such as  $\tilde{P}$ ,  $\tilde{T}$  and  $\tilde{V}$  which are found in the REML equations, (2.8) - (2.10), or in some of the computing algorithms outlined, the matrices in the MME,  $\tilde{X}$  and  $\tilde{Z}$  and sometimes  $\tilde{R}$  and  $\tilde{D}$ , may have relatively simple structures. Obtaining solutions for  $\tilde{\alpha}$  and  $\tilde{b}$  from the MME and using these as estimates in such equations as (2.20) to (2.24) can offer worthwhile computational advantages. Furthermore, use of the MME is often complementary to other algorithms, such as the Fisher-scoring and Newton-Raphson methods, for estimating REML estimates of the variance components. Iteration of the basic REML equations involves inverting  $\tilde{V}$  on each round of the iteration. The matrix  $\tilde{V}$  is never known and furthermore, being of order  $N$ , is usually very large to invert. Based on earlier studies (see, for example, Henderson, 1973 and Harville,

1975, 1977), Searle (1979b) shows that the equations for REML can also be expressed as:

$$\hat{\sigma}_0^2(r+1) = \frac{\underline{y}'(\underline{y} - \underline{X}\hat{\underline{\alpha}}^{(r)} - \underline{Z}\hat{\underline{b}}^{(r)})}{N - p^*} \quad (2.30)$$

and, for  $i = 1, \dots, c$ , the number of random effects:

$$\hat{\sigma}_i^2(r+1) = \frac{(\hat{\underline{b}}_i' \hat{\underline{b}}_i)^{(r)}}{q_i - \text{tr}(\underline{T}_{ii})^{(r)}} \quad (2.31)$$

where  $\hat{\underline{b}}_i$ , a solution for  $\underline{b}_i$ , may be obtained from the MME (e.g., (2.1)),  $\hat{\underline{\alpha}}$  is the estimate of  $\underline{\alpha}$  defined in (2.11),  $p^*$  is the rank of  $\underline{X}$ ,  $q_i$  is the number of levels in the  $i$ th random effect and  $\underline{T}_{ii}$  is defined in (2.21). Searle (1979b) observes that iterating (2.30) and (2.31) always ensures positive estimates because, as he shows, the numerators are always positive,  $\text{tr}(\underline{T}_{ii})$  is always positive and  $q_i$  is always greater than or equal to  $\text{tr}(\underline{T}_{ii})$  when  $\sigma_i^2$  is greater than 0.

The MME can also be usefully applied to obtain an estimate of  $\underline{T}$ . The algorithm for computing  $\hat{\underline{T}}$  from Henderson's MME is given in the appendix. A similar procedure yields  $\hat{\underline{T}}$  from Harville's MME. The advantage of Harville's MME in this respect is that  $\underline{T}$  is directly obtained as the lower right-hand submatrix of the generalised inverse of  $\underline{C}$ , whereas for Henderson's MME some additional computation is required.

#### 4. Modifications to apply constraints

A major problem with the unconstrained gradient algorithms (outlined in section F.2.) for variance component estimation is that they can all produce negative solutions, which are not permitted by the parameter space. Constraints may be applied in a number

of ways, although procedures such as setting negative solutions to zero after each round, or ignoring these solutions completely, are unsatisfactory. Such procedures, although they have been used (e.g., Hemmerle and Hartley, 1973; Miller, 1973), can be the cause of convergence to stationary points which are not maxima. A preferable approach is to modify the unconstrained algorithm with a technique such as those discussed below (see also, Harville, 1977; Bard, 1974 or Gill and Murray, 1974, for example).

To apply non-negativity constraints for variance component estimation there are a number of algorithms available including penalty techniques which maintain the solution within the parameter space by modifying the objective function by some penalty term after each iteration. There are also the transformation techniques which transform the constrained optimization into an unconstrained problem. Another method is the gradient projection technique which "can be used whenever all the inequality constraints [for example,  $\sigma_0^2 > 0$  or  $\sigma_i^2 \geq 0$ ] are linear constraints (as in computing ML or REML estimates of variance components)" (Harville, 1977), that is, whenever the constraints can be expressed as a linear function of  $\theta$ . With this technique the search direction is modified after each iteration by a set of active constraints. These are defined by Powell (in Gill and Murray, 1974) as "a list of those constraints that are satisfied as equalities during an iteration." Methods for evaluating the active set at each iteration, that is active set strategies, are discussed by Gill and Murray (1974) and Aaby and Dempster (1974).

Harville (1977) states that the unconstrained search direction, denoted by  $\tilde{w}_r$  of (2.14) is then modified, to  $\tilde{W}_r$  say, where:

$$\tilde{W}_r = (\tilde{I} - \tilde{w}_r U_r (U_r' \tilde{w}_r U_r)^{-1} U_r') \tilde{w}_r \quad (2.32)$$

where  $U_r$  is the set of active constraints for the  $r$ th iteration.

Harville (1977) suggests that the gradient projection technique is probably the best of the three methods mentioned in the case of variance component estimation. It is considered "superior to the penalty technique for handling linear constraints, especially when it is suspected that the maximizing value may be located on a boundary." It is also better than the transformation methods which should only be used "in conjunction with algorithms that guarantee at least some increase in the value of the objective function on each iteration" because, for example, when computing REML estimates of variance components it may introduce additional stationary points of the likelihood function.

There are other methods available for imposing constraints such as the modified Newton-Raphson algorithm outlined by Harville (1977) and the improved W-transformation by Hemmerle and Lorens (1976). Harville (1977) observes that "the search for improved optimization algorithms is an ongoing process." Indeed, both the continued development of estimation methods and the derivation of computing algorithms for applying these methods, are important ongoing aspects of variance component estimation.

## II. PARAMETER ESTIMATES FOR THE ROMNEY

### A. Non - Genetic Effects

The phenotypic record of an individual has a genetic and an environmental component, although it is the genetic portion that is important in terms of selection. Non-genetic or environmental effects mask the breeding value (B.V.) of the individual as, although they are not controlled by (or inherited with) the same genes as the trait, they may affect its expression. For breeding purposes it is, therefore, useful to identify the important environmental effects and adjust the records for these to give a more accurate indication of B.V.

A number of studies of environmental effects on liveweight in the Romney (e.g., Ch'ang and Rae, 1961; Ch'ang and Rae, 1970; Baker et al., 1974; Baker et al., 1979; Eikje and Johnson, 1979; Jury et al., 1979) have indicated that date of birth, year of birth, sex and birth-rearing rank of the lamb and age of the dam should be considered. Jury et al. found that between a third and a half of the total within flock variation was controlled by environmental effects. Clarke and Rae (1976) state that from the analysis of large numbers of National Flock Recording Scheme (Sheeplan) flocks it has been found that of the total of 40-45% of the variation in weaning weight that is controlled by environmental effects; 5% is due to year effects, 2% is due to age-of-dam effects, 15% is due to birth-rearing rank effects, 5% due to sex effects and 15% is due to date of birth of the lamb.

Eikje (1979) suggests that hogget fleece weight should be adjusted for sex and, possibly, birth-rearing rank and date of birth also. Estimates of environmental effects on hogget fleeceweight (e.g., Baker et al., 1974) have generally shown sex to be the single most important effect on this

trait. Baker et al. (1979) found sex effects accounted for 16% of the total variation in hogget fleece weight, while the other environmental effects together controlled 3-4% of the total variation.

The importance of the environmental effects does change with time. Maternal and other effects associated with the early environment of the lamb such as the age of the dam, date of birth and birth-rearing rank have a lesser effect as the lamb gets older. For example, Baker et al. (1979) found that the environmental effects, particularly date of birth (11%) and birth-rearing rank (18%) controlled a total of 35% of the variation in lamb fleece weight, but a year later all the environmental effects, except sex, were relatively unimportant. The same trend occurs with liveweight although environmental effects have more influence on hogget liveweight than hogget fleece weight.

The influence of sex on hogget traits found in these studies actually includes a grazing mob effect since the sexes are usually run separately from weaning. Eikje (1978a) notes that a number of factors, such as the possibility of interactions between the effects, or the presence of significant within-flock or within-year variation should be considered when choosing a method of adjustment. Jury et al. (1979) and Eikje and Johnson (1979) concluded that specific within-flock and within-year adjustments for environmental effects could be more accurate in flocks weaning more than 150 lambs; except that overall adjustments for birth-rearing rank are suggested unless more than 1000 lambs are being weaned. Overall adjustments are recommended for smaller flocks.

Sheeplan uses within-flock and within-year deviations from the group means to adjust for sex and to take account of whether a ewe first lambed as a hogget or not (dam class, say), but uses overall adjustments

for the other environmental effects. Liveweight records are adjusted by overall correction factors for the effects of birth-rearing rank, date of birth and age of dam (i.e., age within the older ewe class). The effects of sex and dam class are accounted for by within-flock deviations from the group mean in each year (Clarke and Rae, 1976).

Hogget fleece weight was not initially adjusted for environmental effects except to be expressed as a deviation from the group mean for sex and dam class. Callow and Johnson (1982) note that extensive analysis of data from some Sheeplan flocks has indicated that environmental effects account for 18% of variation in hogget fleece weight of hoggets that were not shorn as lambs; although they only contribute 8% of this variation in hoggets that were shorn as lambs. Accordingly, Sheeplan now uses overall correction factors to adjust hogget fleece weight records for birth-rearing rank and age of dam.

Some of the correction factors used by Sheeplan for liveweight and fleece weight, the traits of interest in this study, are presented in Table 2.1. Adjustments for a specific trait which are made for birth-rearing rank (with singles raised as singles as the base level) and age of dam (with mature dams, 4 years old and older, as the base level) are applied by additive overall correction factors. One year old dams do not have an adjustment factor as hogget dams are accounted for by the within-group deviation. Date of birth is adjusted for by a regression coefficient which has been computed as an overall regression from extensive Sheeplan data. All the correction factors are the same for rams and ewes since sex is adjusted for by deviation from the group mean, following the overall adjustments. The traits for which Sheeplan makes adjustment include weaning weight (WWT), autumn liveweight (ALW), winter liveweight (WLW), spring liveweight (SLW) and hogget fleece

weight (HFW). Table 2.1 does not cover the complete range of adjustments available in Sheeplan (e.g., correction factors for triplets are also used and there is another set of correction factors for the meat breeds) but it includes the adjustments relevant to the present study.

The correction factors used by Sheeplan are the difference in environmental effect between the base level and other levels of the effect which have been estimated from national flock data. The estimates agree well in all cases with the estimates presented in Table 2.2.

Correction factors have traditionally been obtained using least squares procedures. The environmental effects have usually been estimated by assuming a conventional fixed effects model (where all terms are regarded as fixed except for a random, residual term) and applying Ordinary Least Squares (OLS). Under OLS the random, residual effects are assumed to be independently and identically distributed; the additional assumption of normality being required for hypothesis testing.

A more general approach, which can be applied to mixed models is Generalised Least Squares (GLS). For the mixed model defined in (2.1), the GLS estimate of  $\underline{\alpha}$ , the vector of fixed effects is:

$$\hat{\underline{\alpha}} = (\underline{X}'\underline{V}^{-1}\underline{X})^{-1}\underline{X}'\underline{V}^{-1}\underline{y} \quad (2.32)$$

where  $\underline{V}$  is the variance-covariance matrix of the vector of observations  $\underline{y}$ . When  $\underline{V} = I\sigma_e^2$ , the GLS and the OLS estimator for  $\underline{\alpha}$  are the same. Under the assumption of normality, the GLS estimate,  $\hat{\underline{\alpha}}$ , can be shown (see, for example, Searle, 1971a) to be the best, linear, unbiased estimator (BLUE) and the maximum likelihood estimator (MLE) of  $\underline{\alpha}$ . These desirable properties of the GLS estimator make it preferable to the OLS estimator in the mixed model case.

TABLE 2.1: CORRECTION FACTORS USED BY SHEEPLAN FOR DUAL-PURPOSE SHEEP BREEDS

Correction factor <sup>1</sup>	Trait <sup>2</sup> (adjustments given in kg.)				
	WWT	ALW	WLW	SLW	HFW
<u>Birth-rearing rank</u>					
SS - TS	2.0	1.8	1.5	0.4	0.15
SS - TT	4.2	3.1	2.7	2.2	0.26
<u>Age of dam</u>					
mat. - 2 y.o.	1.3	1.9	2.4	2.4	-
mat. - 3 y.o.	0.2	0.4	0.9	1.0	-
<u>Date of birth</u>					
M	0.17	0.10	0.09	0.09	-
F	0.17	0.10	0.09	0.08	-

- <sup>1</sup>.SS = single lamb raised as single  
 TS = twin lamb raised as single  
 TT = twin lamb raised as twin  
 2 y.o. = 2 year old dams  
 3 y.o. = 3 year old dams  
 mat. = mature dams, 4 years old and older  
 M = ram lambs  
 F = ewe lambs

<sup>2</sup>.Abbreviations given in text

**TABLE 2.2:** ESTIMATES OF ENVIRONMENT EFFECTS ON WEANING WEIGHT, HOGGET LIVWEIGHT AND HOGGET FLEECE WEIGHT IN ROMNEY SHEEP

Reference	Trait <sup>1</sup>	Environmental Effects <sup>2</sup> (kg.)					
		Age of dam		Birth-rearing rank		Sex	Regression on birthday
		mat.-2 y.o.	mat.-3 y.o.	SS-TS	SS-TT	M-F	
Ch'ang and Rae (1961) <sup>3</sup>	WWT	1.5	0.8	2.5	4.5	2.3	0.10
Ch'ang and Rae (1970)	WWT	2.0	0.4	1.3	4.2	-	0.12
	HLW	2.4	1.0	0.4	2.2	-	0.08
Baker <u>et al.</u> (1974)	WWT	1.3	0.2	2.5	4.2	1.9	0.12
	HLW	1.2	0.2	1.5	2.1	10.8	0.05
	HFW	0.0	0.0	0.0	0.1	0.5	0.01
Jury <u>et al.</u> (1979) <sup>4</sup>	WWT	1.3	0.3	2.0	4.2	2.1	-

1. WWT = weaning weight, HLW = hogget liveweight, HFW = hogget fleece weight.

2. Abbreviations as for Table 2.1.

3. Average of estimates from 2 flocks, over 9 years.

4. Average of 12 flocks.

## B. Heritabilities

Heritability ( $h^2$ ) is defined as the proportion of the phenotypic variance that is due to the additive genetic effects of the genes.

That is:

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2} \quad (2.33)$$

where  $\sigma_a^2$  is the additive genetic variance and  $\sigma_p^2$  is the phenotypic variance. Heritability is important in animal breeding because it is used to predict a number of parameters such as genetic gain from selection and the relative breeding values of individuals within a population.

Heritability cannot be estimated directly from (2.33), because although  $\hat{\sigma}_p^2$  can be obtained,  $\sigma_a^2$  cannot be estimated since additive genetic values are not observable. However,  $\sigma_a^2$ , and therefore  $h^2$ , can be estimated indirectly using records on related animals. The most common relationships used to estimate  $h^2$  are the correlation between records on paternal half-sibs, and the regression of records of the parent on the offspring's records. Both methods make use of relationships which are readily available in sheep breeding data and which are relatively free of non-additive genetic effects.

The paternal half-sib correlation is completely free of maternal effects and dominance variance and contains less than 1/16 of the epistatic variance. The genetic relationship between paternal half-sibs is  $\frac{1}{4}$ ; thus  $\sigma_a^2$  is estimated as 4 times the between sire variance  $\sigma_s^2$ . Similarly,  $h^2$  is estimated as 4 times the paternal half-sib intraclass correlation, which can be expressed as:

$$\hat{h}^2 = 4 \left( \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\sigma}_e^2} \right) \quad (2.34)$$

with  $\sigma_s^2$  denoting the between sire variance and  $\sigma_e^2$  the residual variance.

The paternal half-sib correlation is frequently based on analyses where sires are nested within years; hence, if any sire x year interactions are present then they would be included in  $\sigma_s^2$  of (2.34) (Chopra, 1978). Denoting  $\sigma_{st}^2$  as the variance due to the sire x year interactions, Chopra (1978) suggests the following expression may estimate  $h^2$  more accurately from paternal half-sib records:

$$\hat{h}^2 = 4 \left( \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\sigma}_{st}^2 + \hat{\sigma}_e^2} \right) \quad (2.35)$$

Interactions between fixed and random effects (such as  $\sigma_{st}^2$ ), are frequently ignored, however, because they are difficult to interpret and often non-significant anyway. Nevertheless, if such interactions are significant, but not taken into account and removed from  $\hat{\sigma}_s^2$ , the  $h^2$  will be over-estimated to some extent.

The relationship between parent and offspring can also be used to estimate  $h^2$ . Unlike the parent-offspring correlation, the parent-offspring regression is not biased by selection having been practiced among the parents, the independent variable, which makes it particularly suitable in an animal breeding context. Selection among the offspring biases both the correlation and the regression, but they have the same expectation when selection has not been practiced in either generation. The parent-offspring correlation is free of dominance effects and includes less than  $\frac{1}{4}$  of the epistatic variance. It does however have the disadvantage that maternal effects may inflate the covariance and

therefore the  $h^2$  estimate. Because the genetic relationship between parent and offspring is  $\frac{1}{2}$ , the regression coefficient is doubled to estimate  $h^2$ :

$$\hat{h}^2 = 2 \left( \frac{\hat{\sigma}_{P_{o,d}}}{\hat{\sigma}_{P_d}^2} \right) \quad (2.36)$$

where  $\sigma_{P_{o,d}}$  denotes the covariance between pairs of records on each parent and their offspring and  $\sigma_{P_d}^2$  is the phenotypic variance of the parents' records.

The  $h^2$  estimates in Table 2.3 for a range of traits in the Romney breed are estimated from either a parent-offspring regression (POR) or paternal half-sib (PHS) approach. In some cases the dam-offspring regression (DOR), sire-offspring regression (SOR) or daughter-dam regression (DDR) are used which are specific applications of the POR. Chopra (1978) uses two variations of the PHS. His PHS<sup>1</sup> approach is given in (2.35) and PHS<sup>2</sup> method estimates  $h^2$  as:

$$\hat{h}^2 = 4 \left( \frac{\hat{\sigma}_s^2 + \hat{\sigma}_{st}^2}{\hat{\sigma}_s^2 + \hat{\sigma}_{st}^2 + \hat{\sigma}_e^2} \right) \quad (2.37)$$

and is therefore equivalent to the usual PHS estimate of (2.34). Chopra estimated  $h^2$  from flocks under two different stocking rates (control and intensive) and also from a combination of these records. For the combined stocking rate analysis he also included a term for the sire x stocking rate interaction variance,  $\hat{\sigma}_{st}^2$ , alongside  $\hat{\sigma}_s^2$  in (2.35) and (2.37). There was no significant difference found between the estimates of  $h^2$  from the different stocking rates for the traits reviewed here, so just the range is given in Table 2.3.

Chopra (1978) did, however, find some significant differences

TABLE 2.3: HERITABILITY ESTIMATES FOR VARIOUS LIVEWEIGHT AND FLEECE CHARACTERISTICS IN ROMNEY SHEEP

Estimate	Method	Age	Reference	Comments
<u>GREASY FLEECE WEIGHT</u>				
0.26	PHS	14 mth	McMahon (1943)	most probable: 0.01-0.15
0.04	DDR	14 mth	" "	
0.20	DOR	14 mth	Rae (1946)	
0.35	DOR	14 mth	" "	
0.01-0.15	DOR	14 mth	" (1948)	
0.46	DOR	14 mth	Wright and Stevens (1953)	
0.11	DOR	14 mth	Rae (1958)	
0.31	PHS	14 mth	" "	
0.32	DOR	14 mth	" "	
0.43	PHS	14 mth	Tripathy (1966)	
0.23	PHS	14 mth	Lundie (1971)	
0.002	PHS	-	Radomska and Tyszka (1972)	1st shear
0.324	PHS	-	" " "	2nd shear
0.367	PHS	-	" " "	3rd shear
0.29	PHS	15 mth	Baker <u>et al.</u> (1974)	
0.57	PHS	15 mth	" " "	
0.068	-	-	Radomska and Klewiec (1975)	
0.105	-	-	Radomska and Klewiec (1976)	1st shear
0.225	PHS	14 mth	Rae (1977)	
0.38-0.61	PHS <sup>1</sup>	14 mth	Chopra (1978)	range of stock- ing rates
0.38-0.61	PHS <sup>2</sup>	14 mth	" "	
0.28-0.55	PHS	13 mth	Eikje (1978b)	

47 15  
300 180

TABLE 2.3: (continued)

Estimate	Method	Age	Reference	Comments
<u>GREASY FLEECE WEIGHT</u>				
0.37	PHS	13 mth	Eikje (1978b)	ram hoggets
0.43	PHS	13 mth	" "	ewe hoggets
0.34	PHS	13 mth	" "	singles
0.37	PHS	13 mth	" "	twins
0.12	PHS	4 mth	Baker <u>et al.</u> (1979)	ram lambs
0.21	PHS	4 mth	" " "	ewe lambs
0.15	PHS	4 mth	" " "	
0.27	PHS	13 mth	" " "	ram hoggets
0.41	PHS	13 mth	" " "	ewe hoggets
0.29	PHS	13 mth	" " "	
0.34	DOR	13 mth	" " "	
0.37	SOR	13 mth	" " "	
0.29	-	14 mth	Baker (1977)	
0.28-0.34	PHS	13 mth	Blair (1981)	ewe hoggets
0.07-0.15	PHS	13 mth	" "	ram hoggets
0.14-0.68	PHS	ave. life	" "	ewes
<u>QUALITY NUMBER</u>				
0.18	PHS	14 mth	McMahon (1943)	most probable: 0.35-0.40
0.33	DOR	14 mth	" "	
0.37	DOR	14 mth	Rae (1946)	
0.41	DOR	14 mth	" (1948)	
0.27	DOR	14 mth	" (1950)	
0.25	DOR	mature	" (1958)	
0.27	DOR	14 mth	" "	

TABLE 2.3: (continued)

Estimate	Method	Age	Reference	Comments
<u>QUALITY NUMBER</u>				
0.34	DOR	14 mth	Rae (1958)	
0.47	PHS	14 mth	" "	
0.46-0.59	PHS <sup>1</sup>	14 mth	Chopra (1978)	
0.46-0.72	PHS <sup>2</sup>	14 mth	" "	
0.31-0.39	PHS	13 mth	Blair (1981)	ewe hoggets
0.46-0.58	PHS	13 mth	" "	ram hoggets
0.22-0.73	PHS	ave. life	" "	ewes
<u>STAPLE LENGTH</u>				
0.16	DOR	14 mth	Rae (1946)	
0.19	DOR	14 mth	" "	
0.21	DOR	14 mth	" (1948)	
0.35	DOR	14 mth	" (1950)	
0.29	DOR	mature	" (1958)	
0.35	DOR	14 mth	" "	
0.50	DOR	14 mth	" "	
0.48	PHS	14 mth	" "	1st shear
0.46	DDR	14 mth	Tripathy (1966)	
0.127	-	-	Radomska and Klewiec (1975)	
0.109	-	-	Radomska and Klewiec (1976)	
0.54-0.57	PHS <sup>1</sup>	14 mth	Chopra (1978)	
0.56-0.63	PHS <sup>2</sup>	14 mth	" "	
0.09-0.33	PHS	13 mth	Blair (1981)	ewe hoggets
0.39-0.47	PHS	13 mth	" "	ram hoggets

TABLE 2.3: (continued)

Estimate	Method	Age	Reference	Comments
<u>STAPLE LENGTH</u>				
0.18-0.31	PHS	ave. life	Blair (1981)	ewes
<u>CHARACTER</u>				
0.18	PHS	14 mth	McMahon (1943)	
0.12	DOR	14 mth	" "	
0.00	DOR	14 mth	Rae (1946)	
0.15	DOR	14 mth	" "	
0.16	DOR	14 mth	" "	
0.27	DOR	14 mth	" (1948)	side
0.20	DOR	14 mth	" "	forequarter
0.38	DOR	14 mth	" "	hindquarter
0.25	DOR	mature	" (1958)	
0.16	DOR	14 mth	" "	
0.22	DOR	14 mth	" "	
0.12	PHS	14 mth	" "	
0.24-0.31	PHS <sup>1</sup>	14 mth	Chopra (1978)	
0.28-0.34	PHS <sup>2</sup>	14 mth	" "	
0.25-0.50	PHS	13 mth	Blair (1981)	ewe hoggets
0.27-0.50	PHS	13 mth	" "	ram hoggets
0.18-0.57	PHS	ave. life	" "	ewes
<u>FLEECE COLOUR</u>				
0.22-0.34	PHS <sup>1</sup>	14 mth	Chopra (1978)	greasy
0.31-0.44	PHS <sup>2</sup>	14 mth	" "	greasy
0.10-0.19	PHS <sup>1</sup>	14 mth	" "	scoured
0.30-0.39	PHS <sup>2</sup>	14 mth	" "	scoured

TABLE 2.3: (continued)

Estimate	Method	Age	Reference	Comments
<u>BIRTHWEIGHT</u>				
0.013	-	-	Radomska and Klewiec (1975)	
0.19-0.32	PHS	-	Blair (1981)	ewe lambs
0.32	PHS	-	" "	ram lambs
<u>WEANING WEIGHT</u>				
0.35	DOR	3 mth	Ch'ang and Rae (1961)	
0.30	PHS	3 mth	" " " (1970)	
0.23	DOR	3 mth	" " " "	
0.35	PHS	3 mth	Lundie (1971)	
0.612	PHS	-	Radomska and Tyszka (1972)	
0.18	PHS	3 mth	Baker <u>et al.</u> (1974)	
-0.05	PHS	3 mth	" " "	
0.18	-	-	Radomska and Klewiec (1975)	
0.07-0.18	PHS	-	Eikje (1978b)	mean = 0.14
0.12	PHS	3 mth	" "	ram lambs
0.16	PHS	3 mth	" "	ewe lambs
0.15	PHS	3 mth	" "	singles
0.11	PHS	3 mth	" "	twins
-0.06	PHS	3 mth	Baker <u>et al.</u> (1979)	ram lambs
0.20	PHS	3 mth	" " "	ewe lambs
0.08	PHS	3 mth	Eikje (1978b)	
0.22	DOR	3 mth	" "	
0.08	SOR	3 mth	" "	

TABLE 2.3: (continued)

Estimate	Method	Age	Reference	Comments
<u>WEANING WEIGHT</u>				
0.08	PHS	-	Baker (1977)	
0.24-0.35	PHS	3 mth	Blair (1981)	ewe lambs
0.19	PHS	3 mth	" "	ram lambs
<u>HOGGET (YEARLING) LIVEWEIGHT</u>				
0.46	DDR	14 mth	Tripathy (1966)	
0.51	PHS	13 mth	Ch'ang and Rae (1970)	
0.46	DOR	13 mth	" " " "	
0.015	PHS	12 mth	Radomska and Tyszka (1972)	
0.22	PHS	15 mth	Baker <u>et al.</u> (1974)	
0.22	PHS	15 mth	" " "	
0.107	-	12 mth	Radomska and Klewicz (1975)	
0.21-0.46	PHS <sup>1</sup>	14 mth	Chopra (1978)	
0.47-0.72	PHS <sup>2</sup>	14 mth	" "	
0.23	PHS	13 mth	Baker <u>et al.</u> (1979)	ram hoggets
0.31	PHS	13 mth	" " "	ewe hoggets
0.27	PHS	13 mth	Chopra (1978)	
0.26	DOR	13 mth	" "	
0.23	SOR	13 mth	Baker <u>et al.</u> (1979)	
0.27	PHS	-	Baker (1977)	
0.06-0.52	PHS	13 mth	Blair (1981)	ewe hoggets
0.26-0.42	PHS	13 mth	" "	ram hoggets
<u>LIVEWEIGHT AT OTHER AGES</u>				
0.45	PHS	6 mth	Ch'ang and Rae (1970)	

TABLE 2.3: (continued)

Estimate	Method	Age	Reference	Comments
<u>LIVEWEIGHT AT OTHER AGES</u>				
0.35	DOR	6 mth	Ch'ang and Rae (1970)	
0.39	PHS	9 mth	" " " "	
0.42	DOR	9 mth	" " " "	
0.142	PHS	10 mth	Radomska and Tyszka (1972)	
0.24	PHS	6 mth	Baker <u>et al.</u> (1974)	
0.24	PHS	6 mth	" " "	
0.38	PHS	8 mth	" " "	
0.32	PHS	8 mth	" " "	
0.158	-	10 mth	Radomska and Klewiec (1976)	
0.3-0.41	-	0-3 mth	Ugalde (1978)	range only given
0.22	PHS	5 mth	Baker <u>et al.</u> (1979)	
0.33	DOR	5 mth	" " "	
0.20	SOR	5 mth	" " "	
0.22	PHS	7 mth	" " "	
0.28	DOR	7 mth	" " "	
0.09	SOR	7 mth	" " "	
0.34	PHS	10 mth	" " "	
0.28	DOR	10 mth	" " "	
0.29	SOR	10 mth	" " "	
0.23	PHS	16 mth	" " "	
0.47	DOR	16 mth	" " "	
0.22	SOR	16 mth	" " "	

TABLE 2.3: (continued)

Estimate	Method	Age	Reference	Comments
<u>LIVEWEIGHT AT OTHER AGES</u>				
0.46	PHS	10 mth	Baker (1977)	rams
0.21	PHS	10 mth	" "	ewes
0.26-0.48	PHS	ave. life	Blair (1981)	ewes

between the  $PHS^1$  and  $PHS^2$  estimates of  $h^2$ , which indicates that the interactions are important for that trait. He found sire x year interactions to be highly significant for hogget liveweight and greasy fleeceweight, but non-significant for all the other traits reviewed here. Sire x stocking interactions were highly significant for hogget liveweight and moderately significant for greasy colour grade. In the Romney, apart from reproductive rate, hogget liveweight and greasy fleece weight are the most important selection criteria. The implications of these interactions, particularly for hogget liveweight, suggest that selecting breeding animals under low stocking rates (as may be found in studs), which enhance favourable expression of any genetic potential, is likely to be less effective than selection under higher (commercial) stocking rates. Furthermore, using (2.37), or (2.34), to estimate  $h^2$  for these traits will tend to over-estimate  $h^2$  relative to (2.35).

Blair (1981) studied genetic parameters in a control flock of Romneys, a flock selected for open-face and another selected for fleece weight. These were generally in agreement with each other for the traits reviewed here, so again the range only is given. In other cases, a range of  $\hat{h}^2$  in the table represents a number of replicates. McMahon (1943) presented the  $\hat{h}^2$  range that he considered to be most probable as well as his own estimates.

As a general guide, the magnitude of the heritabilities can be interpreted as low between 0.0 and 0.1, intermediate between 0.1 and 0.3 and high between 0.3 and 1.0 (Turner and Young, 1969).

Most estimates of  $h^2$  in Table 2.3 are made at a particular age, frequently as hoggets (13-15 months), although Blair (1981), for

example, also estimated  $h^2$  from the average lifetime performance records of a ewe flock. When  $h^2$  is estimated within a single sex this is indicated unless it is otherwise obvious (e.g., with DDR estimates).

Blair (1981) found differences in the  $h^2$  estimates between the ram and ewe hoggets although these were generally non-significant. However,  $h^2$  estimates for greasy fleece weight were significantly lower for the ram hoggets. Blair discusses the implications of this in an industry where rams have a greater influence on genetic change, but where genetic parameters tend to be estimated from ewe data. He suggests that response to selection for greasy fleece weight could be less than would otherwise be expected. Baker et al. (1979) found when analysing the sexes separately that the  $h^2$  estimates for liveweight and fleeceweight traits were consistently lower for the ram progeny (not all their  $h^2$  of liveweight estimates for each sex separately are included in Table 2.3). This trend was also supported by Eikje (1978b) for weaning weight and hogget fleeceweight. He points out that this suggests it may be desirable to weigh ram and ewe records differently when predicting breeding values for selection purposes. The estimation of  $h^2$  for each sex separately has not been investigated extensively but these three studies suggest the differences could be important particularly in liveweight and fleeceweight traits which are fundamental selection criteria in a dual purpose breed such as the Romney.

Baker et al. (1979) also investigated the effect of birth-rearing rank on the  $h^2$  of weaning weight. The paternal half-sib estimates were 0.12 for singles and 0.05 for twins. Eikje (1978b) found a similar trend for weaning weight, but not for fleeceweight as his results in Table 2.3 show. Work in other breeds also indicates that

$h^2$  is higher in singles than twins for weaning weight. Gjedrem (1967) has suggested this may be associated with the better level of nutrition that the dam can provide for a single lamb than for each twin. He suggests that the two cases may be considered as estimating  $h^2$  in two different environments. Baker et al. (1979) found that the effect of birth-rearing rank on  $h^2$  declined after weaning, although they did not publish the estimates, which agrees with this theory.

Probably the most noticeable feature of the estimates in Table 2.3 is the wide variation both between and within the various traits. This is true of  $h^2$  estimates in all breeds, and Turner and Young (1969) comment that at least part of the large variation in values is due to the inclusion of estimates with large sampling errors. Some of this variation is due to environmental differences such as age, sex and birth-rearing rank and the method, if any, used to adjust for these factors. There is also, of course, the differences in methods used to estimate  $h^2$  since these may involve different variance components. Generally, however, the production characteristics tend to have moderate to high  $h^2$ 's in both the wool and liveweight traits.

Greasy fleece weight, quality number and staple length, amongst the traits reviewed here, all appear to be reasonably heritable traits in the Romney. The estimates on fleece colour are too few and too variable to be conclusive and character has a low heritability. Most of the  $h^2$  estimates for the wool traits have been made on hogget data. In particular there is little information on important traits such as fleece weight at different ages. The  $h^2$  estimates for greasy lamb fleece weight calculated by Baker et al. (1979) suggest that it is markedly lower than that of hogget fleece weight. Studies in other

breeds (see, for example, Young et al., 1960; Mullaney et al., 1970 and Nawara and Duniec, 1972) have also tended to suggest that  $h^2$  of many production traits increases with age, although the results have not always supported this trend. Nevertheless,  $h^2$  is often sufficiently large, even at early ages, to support worthwhile selection responses in the production wool traits following selection as hoggets or even earlier.

Liveweight also appears to show increased  $h^2$  with increasing age. However, although the estimates of  $h^2$  of birthweight are generally lower than at older ages, the lack of information on this trait makes definite conclusions impossible. Blair (1981) obtained a moderate estimate of  $h^2$  of birthweight in ram lambs from the fleeceweight flock. There was no estimate published for the facecover flock for this trait, due to a negative sire component, and no estimates calculated from the control flock data for any trait in the ram progeny, due to a lack of sufficient numbers. Baker et al. (1979) found that  $h^2$  of liveweight increased with age from weaning to hogget stage. Other estimates in the table at intermediate stages are in general agreement with this despite some obvious exceptions. Baker et al. discuss the importance of recognizing that different methods include different sources of variation. The fact that some maternal effects are included in the numerator of the dam offspring regression as well as the covariance between individual and maternal effects is highly relevant in the case of liveweight traits. The sire offspring regression will include the covariance term but not the variance due to any maternal effects. It is therefore not surprising that  $h^2$  estimates from the dam offspring regression are noticeably higher at weaning and weights close to weaning and that the difference is reduced

with increasing age. The effect of the maternal environmental decreases from an important role in lamb growth to a minimal level in the older lamb and appears to have been eliminated by the hogget stage. Liveweight, at all ages, is also influenced by the individual lamb's own genotype. However, the distinction between and relative magnitude of maternal and individual effects on liveweight from birth onwards has not been fully studied. Eikje (1978b) notes that these two effects on weaning weight should be treated as separate traits. Weaning weight is usually treated as a lamb trait, not a maternal trait. As Baker et al. (1979) observe "there are few estimates in sheep of the heritability of maternal effects or of the genetic correlation between individual and maternal effects", a topic which warrants further study.

In general the liveweight traits, like the production wool traits, do have moderate to high  $h^2$ 's and can also be expected to provide reasonable selection response.

### C. Correlations

The correlation coefficient ( $r$ ) is a measure of the association between two variates. Inter-trait correlations which usually refer to the correlation between two traits measured at the same time on the same animal will be discussed here, although intra-trait correlations on the same or different animals also provide a useful indication of expected changes of correlated variates in a selected population. The phenotypic correlation may be divided into a genetic and an environmental portion. It is the genetic correlation that determines the correlated response that may be expected from selection. Phenotypic correlations play an important role in constructing selection indices.

The phenotypic correlation between two traits ( $r_p$ ) may be

estimated from pairs of records on a sample of individuals from a population, using the product-moment expression. That is,

$$\hat{r}_p = \frac{\hat{\sigma}_{P_{1,2}}}{\sqrt{\hat{\sigma}_{P_1}^2 \cdot \hat{\sigma}_{P_2}^2}} \quad (2.38)$$

where  $\sigma_{P_{1,2}}$  denotes the covariance between two traits 1 and 2 calculated from pairs of records from a number of individuals, and  $\sigma_{P_1}^2$  and  $\sigma_{P_2}^2$  represent the phenotypic variances of the two traits. When variance and covariance components have been estimated from the data then  $r_p$  can readily be estimated. Thus, following an analysis of variance and covariance from paternal half-sib data:

$$\hat{r}_p = \frac{\hat{\sigma}_{s_{1,2}} + \hat{\sigma}_{e_{1,2}}}{\sqrt{(\hat{\sigma}_{s_1}^2 + \hat{\sigma}_{e_1}^2)(\hat{\sigma}_{s_2}^2 + \hat{\sigma}_{e_2}^2)}} \quad (2.39)$$

where  $\sigma_s^2$  and  $\sigma_e^2$  denote the sire and residual variance components, and  $\sigma_{s_{1,2}}$  and  $\sigma_{e_{1,2}}$  the sire and residual covariance terms respectively. The covariance may be of any magnitude and sign, and therefore the correlation can also be positive or negative. However, the standardisation afforded by the denominator, theoretically ensures that the correlation ranges between -1 and +1. Generally, an absolute value between 0 and 0.2 is regarded as negligible, 0.2 to 0.4 as low, 0.4 to 0.6 as moderate and over 0.6 as high (Brown and Turner, 1968). These guidelines apply to both phenotypic and genetic correlations.

Estimation of the genetic correlations is not as straightforward as the estimation of the phenotypic correlation using (2.38). Once again the paternal half-sib and parent-offspring relationships are most frequently used. The necessary variance and covariance components can

be estimated from the data with appropriate analyses of variance and covariance. From the paternal half-sib analysis, the covariance between the traits due to sires estimates a quarter of the additive genetic variance, in the same way that  $\sigma_s^2$  estimates  $\frac{1}{4}\sigma_a^2$ , as was discussed for heritability. Thus:

$$\hat{r}_g = \frac{\hat{\sigma}_{s_{1,2}}}{\sqrt{\hat{\sigma}_{s_1}^2 \cdot \hat{\sigma}_{s_2}^2}} \quad (2.40)$$

since the fours cancel out. The genetic correlation can also be estimated from the parent-offspring regression although the expression is somewhat different in that it involves 2 different traits and 2 classes of individuals on which measurements are made. That is

$$r_g = \frac{\hat{\sigma}_{g_{d_1 o_2}} + \hat{\sigma}_{g_{d_2 o_1}}}{2\sqrt{\hat{\sigma}_{g_{d_1 o_1}} \cdot \hat{\sigma}_{g_{d_2 o_2}}}} \quad (2.41)$$

The 2 is introduced in the denominator to obtain the mean of the 2 covariances in the numerator. In the denominator the sums of cross products although calculated within traits, are between animals thus covariances rather than variances are employed.

In part I a number of the methods available for estimating variance components were discussed. These methods can generally be applied to the estimation of covariance components also (see Henderson, 1953, for example).

The apportioning of the phenotypic correlation into a genetic and an environmental component is illustrated in equation (2.39). A positive correlation between two traits does not however imply they have a common genetic base. The genetic component could be small and/or negative and it could be the environmental component which is

causing the positive value for  $\sigma_p^2$ . The reverse could also apply, but  $\sigma_p^2$  itself does not convey this information.

The genetic correlation also indicates what, if any, counter-selection may be necessary. It is also possible, when a desired characteristic is difficult or expensive to measure, that there may be a highly genetically correlated trait on which indirect selection can be practiced. This also applies to phenotypic correlations with respect to lifetime performance. It is therefore evident that the selection objectives could be antagonistic if the genetic and phenotypic correlations have different signs. Since the phenotypic correlation involves both a genetic and an environmental component, whereas the genetic correlation just the former, they are more likely to be of the same sign for traits in which the environment has a lesser effect. Pleiotropy, and to some extent linkage, are important causes of the genetic correlation.

In Tables 2.4 and 2.5 a range of estimates of phenotypic correlations and genetic correlations, respectively, are presented. As with the heritabilities, the estimates reviewed are confined to a selection of wool and growth traits that are of interest in the Romney breed. The phenotypic correlation is invariably estimated from the product-moment expression of (2.38); hence the method of estimation is not listed. In both tables, age pertains to both traits in the correlation, unless it is indicated to be otherwise. The fleece trait x liveweight correlations are exceptions since they generally refer to characteristics of the hogget fleece, with the liveweights being recorded at the ages specified. The range of correlations presented from Blair (1981) cover the same selection flocks described

TABLE 2.4: PHENOTYPIC CORRELATION ESTIMATES BETWEEN VARIOUS LIVE-WEIGHT AND FLEECE CHARACTERISTICS IN ROMNEY SHEEP

Estimate	Age	Reference	Comments
<u>GREASY FLEECE WEIGHT x HOGGET FLEECE WEIGHT</u>			
0.51	2 y.o.	Wright and Stevens (1953)	
0.35	5 y.o.	" " " "	
0.44	ave. life	" " " "	excl. hoggets
0.62	ave. life	" " " "	incl. hoggets
<u>GREASY FLEECE WEIGHT x QUALITY NUMBER</u>			
-0.33	14 mth	Rae (1958)	
-0.07--0.03	14 mth	Sumner (1969)	
-0.31--0.18	mixed	" "	ewes
-0.33	mixed	Wickham <u>et al.</u> (unpubl.)	ewes
-0.18	14 mth	Eikje (1978b)	ewe hoggets
-0.13	14 mth	" "	ram hoggets
-0.27	14 mth	" "	both sexes
-0.26--0.22	13 mth	Blair (1981)	ewe hoggets
-0.26--0.05	13 mth	" "	ram hoggets
<u>GREASY FLEECE WEIGHT x STAPLE LENGTH</u>			
0.45	14 mth	Rae (1958)	
0.48	14 mth	Tripathy (1966)	
0.22-0.51	14 mth	Sumner (1969)	
0.30-0.39	mixed	" "	ewes
0.47	mixed	Wickham <u>et al.</u> (unpubl.)	ewes
0.17	-	Radomska and Klewicz (1975)	
0.20	-	" " " (1976)	
0.40-0.50	13 mth	Blair (1981)	ewe hoggets
0.29-0.51	13 mth	" "	ram hoggets

TABLE 2.4: (continued)

Estimate	Age	Reference	Comments
<u>GREASY FLEECE WEIGHT x CHARACTER</u>			
0.24	14 mth	Rae (1958)	
0.12-0.22	14 mth	Sumner (1969)	
-0.01-0.18	mixed	" "	ewes
-0.09	mixed	Wickham <u>et al.</u> (unpubl.)	ewes
0.06-0.07	13 mth	Blair (1981)	ewe hoggets
0.00-0.18	13 mth	" "	ram hoggets
<u>QUALITY NUMBER x STAPLE LENGTH</u>			
-0.54	14 mth	Rae (1958)	
-0.69	mixed	Ross (1964)	ewes
-0.49--0.46	14 mth	Sumner (1969)	
-0.52--0.42	mixed	" "	ewes
-0.53	mixed	Wickham <u>et al.</u> (unpubl.)	ewes
-0.58--0.53	13 mth	Blair (1981)	ewe hoggets
-0.55--0.31	13 mth	" "	ram hoggets
<u>QUALITY NUMBER x CHARACTER</u>			
0.06	14 mth	Rae (1958)	
0.39-0.58	14 mth	Sumner (1969)	
0.08-0.38	mixed	" "	ewes
0.33	mixed	Wickham <u>et al.</u> (unpubl.)	ewes
1.11-1.13	13 mth	Blair (1981)	ewe hoggets
0.13-0.36	13 mth	" "	ram hoggets
<u>STAPLE LENGTH x CHARACTER</u>			
0.20	14 mth	Rae (1958)	
-0.26--0.11	14 mth	Sumner (1969)	

TABLE 2.4: (continued)

Estimate	Age	Reference	Comments
<u>STAPLE LENGTH x CHARACTER</u>			
-0.10-0.05	mixed	Sumner (1969)	ewes
-0.27	mixed	Wickham <u>et al.</u> (unpubl.)	ewes
-0.15--0.09	13 mth	Blair (1981)	ewe hoggets
-0.15-0.12	13 mth	" "	ram hoggets
<u>GREASY FLEECE WEIGHT x LIVEWEIGHT</u>			
0.61	14 mth	Tripathy (1966)	
-0.50-0.52	14 mth	Sumner (1969)	
-0.27-0.49	mixed	" "	ewe ave. life FLW
0.47	mixed	Wickham <u>et al.</u> (unpubl.)	ewe ave. life FLW
0.11	birth	Radomska and Klewiec (1975)	
0.17	3 mth	" " " "	
0.23	12 mth	" " " "	
0.20	10 mth	" " " (1976)	
0.29	3 mth	Eikje (1978b)	all progeny
0.30	3 mth	" "	ram progeny
0.26	3 mth	" "	ewe progeny
0.30	3 mth	" "	singles
0.30	3 mth	" "	twins
0.25	3 mth	Baker <u>et al.</u> (1979)	
0.38	5 mth	" " "	
0.52	10 mth	" " "	
0.44	13 mth	" " "	
0.40	16 mth	" " "	

TABLE 2.4: (continued)

Estimate	Age	Reference	Comments
<u>GREASY FLEECE WEIGHT x LIVEWEIGHT</u>			
0.47-0.55	13 mth	Blair (1981)	ewe hoggets
0.44-0.53	13 mth	" "	ram hoggets
<u>QUALITY NUMBER x LIVEWEIGHT</u>			
0.08-0.10	14 mth	Sumner (1969)	
0.04-0.14	mixed	" "	ewe ave. life QN
-0.01	3 mth	Eikje (1978b)	all progeny
-0.01	3 mth	" "	ewe progeny
-0.03	3 mth	" "	ram progeny
-0.08-0.05	13 mth	Blair (1981)	ewe hoggets
-0.07--0.02	13 mth	" "	ram hoggets
<u>STAPLE LENGTH x LIVEWEIGHT</u>			
0.24	14 mth	Tripathy (1966)	
0.01-0.18	14 mth	Sumner (1969)	
0.09-0.17	mixed	" "	ewe ave. life SL
0.14	mixed	Wickham <u>et al.</u> (unpubl.)	ewe ave. life SL
-0.06	birth	Radomska and Klewicz (1975)	
-0.01	3 mth	" " " "	
0.06	12 mth	" " " "	
0.15-0.26	13 mth	Blair (1981)	ewe hoggets
0.03-0.21	13 mth	" "	ram hoggets

TABLE 2.4: (continued)

Estimate	Age	Reference	Comments
<u>BIRTHWEIGHT x LIVEWEIGHT AT LATER AGES</u>			
0.32	3 mth	Radomska and Klewiec (1975)	
0.27	12 mth	" " " "	
<u>WEANING WEIGHT x LIVEWEIGHT AT LATER AGES</u>			
0.82	6 mth	Ch'ang and Rae (1972)	
0.74	9 mth	" " " "	
0.62	13 mth	" " " "	
0.61	12 mth	Radomska and Klewiec (1975)	
0.69	5 mth	Baker <u>et al.</u> (1979)	
0.57	10 mth	" " "	
0.49	13 mth	" " "	
0.45	16 mth	" " "	
<u>HOGGET (YEARLING) WEIGHT BY LIVEWEIGHT AT VARIOUS AGES</u>			
0.74	6 mth	Ch'ang and Rae (1972)	
0.85	9 mth	" " " "	
0.60	5 mth	Baker <u>et al.</u> (1979)	
0.83	10 mth	" " "	
0.70	16 mth	" " "	
<u>BETWEEN LIVEWEIGHTS AT VARIOUS OTHER AGES</u>			
0.89	6 x 9 mth	Ch'ang and Rae (1972)	
0.72	5 x 10 mth	Baker <u>et al.</u> (1979)	
0.57	5 x 16 mth	" " "	
0.70	10 x 16 mth	" " "	

for heritability. The range of estimates from Sumner (1969) represent two different stocking rates, control and intensive.

From Table 2.4 and 2.5 it appears that the genetic and phenotypic correlations generally agree in terms of sign. As mentioned earlier this is an important advantage in terms of selection objectives. The genetic correlation is usually of greater magnitude than the phenotypic correlations. However, as with  $h^2$ , there is considerable variation within each trait. Although this is often associated with high standard errors of the estimates, it may also reflect changes in the genetic variation due to factors such as selection pressure, inbreeding and outbreeding. The phenotypic variation may also change over time due to both genetic and environmental changes. Changes in management policies and nutritional regimes are typical examples. Thus it is desirable to re-estimate parameters involving variance components, such as the correlations and heritabilities, fairly regularly. High standard errors and estimates lying outside the bounds of the correlation have often been associated with the analysis of low numbers of records. However this is not always the case. Several of the genetic correlations generated by Radomska and Klewiec (1975) have a magnitude greater than 1 although 1717 paternal half-sib records were analysed.

In both tables, there does not appear to be a great deal of difference in either the genetic or phenotypic correlations between the sexes. Differences in birth-rearing rank have only been studied for the correlations between greasy fleece weight and weaning weight (Eikje, 1978b). The phenotypic correlation is no different between the singles and twins and the genetic correlation is higher for the twins; however more estimates of these differences are needed to make conclusions.

The estimation of genetic correlations by different methods has been studied by several workers in the table. Baker et al. (1979) comment that for their estimates between the liveweight traits, "with a few exceptions estimates of genetic correlations obtained by different methods of estimation showed satisfactory agreement, although sampling errors were relatively large (0.01-0.27)." Ch'ang and Rae (1972) also estimated genetic correlations between the liveweights by different methods, and their results also appear to support this, although they found a tendency for the DDR estimates to be slightly higher than the PHS estimates. They suggested this could be due to the maternal effects ignored in the DDR and therefore estimated the correlation by a third method using dam-offspring covariance components. These estimates were little different from the DDR estimates and suggested the difference in the PHS estimates was a property of the data, for the between liveweight correlations, rather than inadequate consideration of the maternal effects by the PHS method.

In general it is evident from the table that both the genetic and phenotypic correlations are reasonably high between all the liveweights. The genetic correlations between weaning weight and later ages suggest preliminary selection can be made at an early age.

The phenotypic and genetic correlations between liveweight and fleece weight are positive, although fairly low, which is an important advantage for a dual purpose sheep such as the Romney. The correlations between liveweight and fleece quality traits are considerably more variable. They are generally low, although the phenotypic correlations are particularly variable.

Correlations between the fleece weight and fleece quality traits

are also variable and can be antagonistic to some extent. Thus although greasy fleece weight, clean fleece weight and staple length tend to be positively correlated with each other, they are generally negatively correlated with quality number and character and positively correlated with mean fibre diameter and hairiness. The noticeably low correlations involving character is one of the reasons that this trait has become regarded as a less useful selection objective.

In the Romney, fleece weight and liveweight are two important objectives of improvement. Depending on the end-use demand for the wool either coarser or finer fibre diameter may be preferred. So although fleece weight and liveweight can be selected for together with little difficulty, it may be necessary to monitor changes in fibre diameter so that these can be maintained at the desired level.

TABLE 2.5: GENETIC CORRELATION ESTIMATES BETWEEN VARIOUS LIVEWEIGHT AND FLEECE CHARACTERISTICS IN ROMNEY SHEEP

Estimate	Method	Age	Reference	Comments
<u>GREASY FLEECE WEIGHT x HOGGET FLEECE WEIGHT</u>				
0.65	PHS	2 y.o.	Radomska and Tyszka (1972)	2nd x 3rd shear
0.80	PHS	3 mth	Baker (1977)	
<u>GREASY FLEECE WEIGHT x QUALITY NUMBER</u>				
-0.62	PHS	14 mth	Rae (1958)	
-0.47	DOR	mature	" "	
-0.49	DOR	14 mth	" "	
-0.47	DOR	14 mth	" "	
-0.09	PHS	14 mth	Eikje (1978b)	ewe hoggets
0.88	PHS	14 mth	" "	ram hoggets
-0.18	PHS	14 mth	" "	both sexes
0.02	PHS	14 mth	Chopra (1978)	
-0.43--0.07	PHS	13 mth	Blair (1981)	ewe hoggets
-0.24	PHS	13 mth	" "	ram hoggets
<u>GREASY FLEECE WEIGHT x STAPLE LENGTH</u>				
0.60	PHS	14 mth	Rae (1958)	
0.26	DOR	mature	" "	
0.21	DOR	14 mth	" "	
0.25	DOR	14 mth	" "	
0.40	DDR	14 mth	Tripathy (1966)	
0.54	-	-	Radomska and Klewiec (1975)	
0.75	-	-	Radomska and Klewiec (1976)	
0.58	PHS	14 mth	Chopra (1978)	

TABLE 2.5: (continued)

Estimate	Method	Age	Reference	Comments
<u>GREASY FLEECE WEIGHT x STAPLE LENGTH</u>				
0.35-0.58	PHS	13 mth	Blair (1981)	ewe hoggets
0.44-0.56	PHS	13 mth	" "	ram hoggets
<u>GREASY FLEECE WEIGHT x CHARACTER</u>				
0.27	PHS	14 mth	Rae (1958)	
0.09	DOR	mature	" "	
-0.16	DOR	14 mth	" "	
0.08	DOR	14 mth	" "	
0.24	PHS	14 mth	Chopra (1978)	
-0.17-0.39	PHS	13 mth	Blair (1981)	ewe hoggets
-0.36--0.16	PHS	13 mth	" "	ram hoggets
<u>GREASY FLEECE WEIGHT x COLOUR</u>				
-0.19	PHS	14 mth	Chopra (1978)	greasy
-0.11	PHS	14 mth	" "	scoured
<u>QUALITY NUMBER x STAPLE LENGTH</u>				
-0.76	PHS	14 mth	Rae (1958)	
-0.73	DOR	14 mth	" "	
-0.53	PHS	14 mth	Chopra (1978)	
-0.93--0.37	PHS	13 mth	Blair (1981)	ewe hoggets
-0.65	PHS	13 mth	" "	ram hoggets
<u>QUALITY NUMBER x CHARACTER</u>				
-0.41	PHS	14 mth	Rae (1958)	
0.21	DOR	14 mth	" "	
0.61	PHS	14 mth	Chopra (1978)	

TABLE 2.5: (continued)

Estimate	Method	Age	Reference	Comments
<u>QUALITY NUMBER x CHARACTER</u>				
-0.93--0.37	PHS	13 mth	Blair (1981)	ewe hoggets
0.57-0.64	PHS	13 mth	" "	ram hoggets
<u>QUALITY NUMBER x COLOUR</u>				
0.02	PHS	14 mth	Chopra (1978)	greasy
0.27	PHS	14 mth	" "	scoured
<u>STAPLE LENGTH x CHARACTER</u>				
0.74	PHS	14 mth	Rae (1958)	
0.13	DOR	14 mth	" "	
-0.28	PHS	14 mth	Chopra (1978)	
-0.78--0.06	PHS	13 mth	Blair (1981)	ewe hoggets
-0.45--0.37	PHS	13 mth	" "	ram hoggets
<u>STAPLE LENGTH x COLOUR</u>				
-0.06	PHS	14 mth	Chopra (1978)	greasy
-0.23	PHS	14 mth	" "	scoured
<u>CHARACTER x COLOUR</u>				
0.57	PHS	14 mth	Chopra (1978)	greasy
0.20	PHS	14 mth	" "	scoured
<u>GREASY FLEECE WEIGHT x LIVEWEIGHT</u>				
0.54	PHS	14 mth	Tripathy (1966)	
1.00	PHS	12 mth	Radomska and Tyszka (1972)	3rd shear FWT
-1.29	-	birth	Radomska and Klewiec (1975)	
0.30	-	3 mth	" " "	
0.39	-	12 mth	" " "	

TABLE 2.5: (continued)

Estimate	Method	Age	Reference	Comments
<u>GREASY FLEECE WEIGHT x LIVEWEIGHT</u>				
0.64	-	10 mth	Radomska and Klewiec (1976)	
0.29	PHS	3 mth	Eikje (1978b)	all progeny
0.29	PHS	3 mth	" "	ram progeny
0.37	PHS	3 mth	" "	ewe progeny
0.23	PHS	3 mth	" "	singles
0.60	PHS	3 mth	" "	twins
0.11	PHS	14 mth	Chopra (1978)	
0.39	PHS	3 mth	Baker <u>et al.</u> (1979)	
0.24	DOR	3 mth	" " "	
0.04	SOR	3 mth	" " "	
0.37	PHS	5 mth	" " "	
0.08	DOR	5 mth	" " "	
0.05	SOR	5 mth	" " "	
0.40	PHS	10 mth	" " "	
0.21	DOR	10 mth	" " "	
0.11	SOR	10 mth	" " "	
0.41	PHS	13 mth	" " "	
-0.04	DOR	13 mth	" " "	
0.02	SOR	13 mth	" " "	
0.40	PHS	16 mth	" " "	
0.26	DOR	16 mth	" " "	
-0.14	SOR	16 mth	" " "	
0.64-0.89	PHS	13 mth	Blair (1981)	ewe hoggets
0.49-0.59	PHS	13 mth	" "	ram hoggets

TABLE 2.5: (continued)

Estimate	Method	Age	Reference	Comments
<u>QUALITY NUMBER x LIVEWEIGHT</u>				
-0.29	PHS	3 mth	Eikje (1978b)	all progeny
-0.11	PHS	3 mth	" "	ewe progeny
0.40	PHS	3 mth	" "	ram progeny
0.37	PHS	14 mth	Chopra (1978)	
-0.71-0.12	PHS	13 mth	Blair (1981)	ewe hoggets
-0.18-0.47	PHS	13 mth	" "	ram hoggets
<u>STAPLE LENGTH x LIVEWEIGHT</u>				
0.21	DDR	14 mth	Tripathy (1966)	
2.54	-	birth	Radomska and Klewiec (1975)	
0.44	-	3 mth	" " "	
0.73	-	12 mth	" " "	
0.50	PHS	14 mth	Chopra (1978)	
-0.20-0.77	PHS	13 mth	Blair (1981)	ewe hoggets
0.21-0.47	PHS	13 mth	" "	ram hoggets
<u>BIRTHWEIGHT x LIVEWEIGHT AT LATER AGES</u>				
0.56	-	3 mth	Radomska and Klewiec (1975)	
1.85	-	12 mth	" " "	
<u>WEANING WEIGHT x LIVEWEIGHT AT LATER AGES</u>				
0.73	PHS	6 mth	Ch'ang and Rae (1972)	
0.96	DDR	6 mth	" " " "	
0.86	PHS	9 mth	" " " "	
0.97	DDR	9 mth	" " " "	
0.74	PHS	13 mth	" " " "	

TABLE 2.5: (continued)

Estimate	Method	Age	Reference	Comments
<u>WEANING WEIGHT x LIVEWEIGHT AT LATER AGES</u>				
0.90	DDR	13 mth	Ch'ang and Rae (1972)	
0.86	-	12 mth	Radomska and Klewiec (1975)	
0.78	PHS	5 mth	Baker <u>et al.</u> (1979)	
1.02	DOR	5 mth	" " "	
0.80	SOR	5 mth	" " "	
0.74	PHS	10 mth	" " "	
0.72	DOR	10 mth	" " "	
0.70	SOR	10 mth	" " "	
0.77	PHS	13 mth	" " "	
0.62	DOR	13 mth	" " "	
0.76	SOR	13 mth	" " "	
0.73	PHS	16 mth	" " "	
0.60	DOR	16 mth	" " "	
0.56	SOR	16 mth	" " "	
<u>HOGGET (YEARLING) WEIGHT x LIVEWEIGHT AT VARIOUS AGES</u>				
0.86	PHS	6 mth	Ch'ang and Rae (1972)	
0.86	DDR	6 mth	" " " "	
0.96	PHS	9 mth	" " " "	
0.90	DDR	9 mth	" " " "	
0.87	PHS	5 mth	Baker <u>et al.</u> (1979)	
0.83	DOR	5 mth	" " "	
1.03	SOR	5 mth	" " "	
0.97	PHS	10 mth	" " "	

TABLE 2.5: (continued)

Estimate	Method	Age	Reference	Comments
<u>HOGGET (YEARLING) WEIGHT x LIVEWEIGHT AT VARIOUS AGES</u>				
0.90	DOR	10 mth	Baker <u>et al.</u> (1979)	
0.95	SOR	10 mth	" " "	
0.98	PHS	16 mth	" " "	
0.98	DOR	16 mth	" " "	
1.00	SOR	16 mth	" " "	
<u>BETWEEN LIVEWEIGHTS AT VARIOUS OTHER AGES</u>				
0.96	PHS	6 x 9 mth	Ch'ang and Rae (1972)	
0.98	DDR	6 x 9 mth	" " " "	
0.87	PHS	5 x 10 mth	Baker <u>et al.</u> (1979)	
1.05	DOR	5 x 10 mth	" " "	
0.91	SOR	5 x 10 mth	" " "	
0.84	PHS	5 x 16 mth	" " "	
0.85	DOR	5 x 16 mth	" " "	
0.94	SOR	5 x 16 mth	" " "	
0.96	PHS	10 x 16 mth	" " "	
0.99	DOR	10 x 16 mth	" " "	
0.95	SOR	10 x 16 mth	" " "	

CHAPTER THREE

SOURCE OF DATA

## CHAPTER THREE

SOURCE OF DATAI. RUAKURA SHEEP SELECTION STUDIESA. Introduction and Objectives

This study is concerned with various liveweight and fleece traits recorded on the Romney lambs born from 1970 to 1972 at Woodlands Research Station near Invercargill. This represents a section of data from the establishment phase of a long-term selection experiment being undertaken by the Genetics Section of Ruakura Animal Research Station at Hamilton. A brief outline is given of the full experiment, before discussing details of the data set specific to the present analysis, in part II. The information has been assembled from a number of sources including the Genetics Section of the MAF Annual Reports, in particular the 1970/71, 1971/72 and 1972/73 issues, several unpublished publications from Ruakura Genetics Section, and Clarke (pers. comm.) and Hickey (pers. comm.).

The Ruakura sheep selection studies commenced in 1970 and have an intended duration of 15 to 20 years. The objectives are to provide estimates of direct and correlated responses to selection for reproductive, growth and wool production traits, and estimates of various genetic and phenotypic parameters. It is also intended, where possible, to determine the biological and physiological bases of the genetic responses to selection. Estimates of the genetic and phenotypic parameters will be available for use in the national flock recording scheme, Sheeplan.

There are a number of different groups within which selection is being practised. These include 3 different locations and 7 different breed groups although breeds are neither completely nested within, nor cross-classified across locations. There are 6 breeds in the selection study; Romney, Cheviot, Corriedale, Dorset (both horned and polled animals) and Border-Romney cross ewes. The seventh breed group consists of Romneys screened for high fertility with the requirement being that two-tooth ewes have weaned a set of twins and older ewes at least two lambs from a set of triplets. The 3 locations, Tokanui (near Hamilton), Templeton (near Christchurch) and Woodlands (near Invercargill), are all Ministry of Agriculture and Fisheries Research Stations located throughout New Zealand. Hence, the comparative selection response between the different breeds and environments are two further important objectives of the experiment.

## B. Design of the Experiment

### 1. The establishment phase

The Ruakura sheep selection experiment has proceeded in 2 distinct phases. The first establishment phase includes the 1970 to 1972 matings. The screened Romneys were purchased from Southland flocks only during 1970-1973. Ewes in the other breeding groups were purchased from commercial farmers, breed societies and research institutions all over New Zealand between 1970-1972. There were 60 different sources for the Romney, 50 for the Border Leicester, 20 for the Cheviot and 15 for both the Dorset and Corriedale ewes. These were collected and distributed evenly each year between the 3 locations. Each year of the 1970-1973 establishment phase there was an interchange of sheep just prior to mating between the 3

locations; Tokanui, Templeton and Woodlands. This was done to establish a base population with similar genetic composition in all 3 flocks. This interchange also makes it possible to estimate the extent to which performance is effected by different climates, locations and managerial regimes; although management was kept as uniform as possible.

There was no selection practised during the establishment phase. All progeny were retained to hogget stage and all rams were left entire. Numbers were reduced to the required level by a random culling of hoggets across all the flocks prior to re-distribution. The ewes were randomly allocated to sires, which were also purchased from a wide range of breeders throughout New Zealand. The methods of management were standardized as much as possible between the locations, and different classes of sheep were kept in single mobs in an effort to maintain uniform conditions.

## 2. The selection phase

Before mating in 1973 the flocks at Tokanui, Templeton and Woodlands were closed. There was no further interchanging of ewes between the 3 locations, nor were any more sheep brought in from other sources. The ewes were divided into 22 flocks with 9 of these at Tokanui, 5 at Templeton and 8 at Woodlands. There are 15 Romney flocks, 1 screened Romney, 2 Border Leicester, 1 Border-Romney cross, 1 Corriedale, 1 Dorset and 1 Cheviot flock. The sheep were divided into flocks containing 150 mixed-age ewes and some young replacement stock.

Each flock was assigned a selection policy for the duration of the selection phase. The Romney was the only breed with control

flocks. There are 7 different selection policies, including 2 selection indices comprising combinations of the single-trait selection criteria. The details of the selection policies and their allocation across the different breeds and locations are shown in Table 3.1. The total number of flocks at each location and the approximate total number of breeding ewes maintained each year are also shown in the table.

Details of management during the selection phase are not considered in this study.

TABLE 3.1: SHEEP SELECTION EXPERIMENTS AT TOKANUI, TEMPLETON AND WOODLANDS

Selection Criteria	Tokanui			Templeton			Woodlands		
Control	R	R	R	R			R		
Fertility Index	B	R	BR	R					
Number of lambs born	R								
Hogget body weight	R								
Uncommitted	R								
Productivity Index				Co	R	D	B	R	Ch
Number of lambs born							R	SR	
Weaning weight							R		
Hogget fleece weight							R		
Number of flocks	9			5			8		
Number of breeding ewes	1350			750			1200		

Note: R = Romney, B = Border Leicester, BR = Border Romney,

Co = Corriedale, D = Dorset, Ch = Cheviot, SR = Screened

Romney

## II. ESTABLISHMENT PHASE OF THE WOODLANDS ROMNEY FLOCKS

### A. The Sheep and their Environment

During 1969 mixed-age Romney ewes were purchased for the sheep selection study. Early in 1970 a random sample of 762 of these ewes were transferred to Woodlands Research Station. The Woodlands ewes represent 24 different sources of Romney breeders and research institutions throughout New Zealand. There were also 24 Romney sires from 20 different sources sent to Woodlands prior to mating in 1970.

The ewes were randomly allocated, each year, into sire groups with comparable age structures and range of source and liveweight. In 1970, the start of mating was delayed until 10 May as some of the ewes were late arriving at Woodlands. In subsequent years, mating started about 3 weeks earlier. The ewes also took longer to adjust to the new environment in the first year. Furthermore, the weather conditions at lambing were particularly poor in 1970. The mating period lasted 5 weeks in 1970 and 6 weeks in the following 2 years starting on 21 April in 1971 and 17 April in 1973. The ewes were paddock mated in single sire groups.

Before mating in 1971 all the ewes at Tokanui, Templeton and Woodlands, including some additional ewes purchased during the year were randomly re-allocated between the 3 stations and the required interchanges were made. In 1971 there were 741 mixed-age Romney ewes at Woodlands which were again mated in 24 single-sire groups. The rams came from 21, and the ewes from 55 different sources. Rams are generally used as sires for 1 season only, although 2 of the rams used at Woodlands in 1970 were used again in the second year. Mating was slightly delayed due to the exchanging of ewes between the 3 sites.

During the winter of 1971, feed shortages occurred which were accentuated by insufficient supplementary crops.

Prior to mating in the final year of the establishment phase the ewes were once again randomly re-distributed between the 3 stations. In 1972 there were 739 mixed-age Romneys re-allocated to Woodlands. This year the rams included 8 bred at Tokanui and 9 bred at Woodlands with the balance of the 24 sires originating from different Romney breeders. In order to avoid further problems associated with under-nutrition, all hoggets were grazed off the station during the 1972 winter and early spring. The ram hoggets were sent to a station at Te Anau and the ewe hoggets to a Lands and Survey block at Mossburn. Feed was in fact also limited at Te Anau; however, the ram hoggets grew well enough to still be used as sires in 1973. Liveweight recordings were maintained on the hoggets at approximately 2 monthly intervals.

Woodlands Research Station comprises 183 ha of fertile, Southland lowlands (47m above sea level) located 19 km north-east of Invercargill. Temperatures are generally in the range of 5-15°C although -11°C frosts can be expected for a 120-day period over winter. Rainfall is spread evenly throughout the year with a range of 1000-1200 mm. Approximately 1650 hours of sunshine may be expected each year.

The main soil type at Woodlands is the Waikiwi silt loam. The parent material is tuffaceous greywacke over weathered gravels. The original browntop pasture has predominantly been replaced by ryegrass and white clover. The pastures respond well to superphosphate, lime and potash. A stocking rate of 15-18 s.u/ha is carried at present.

## B. The Data Collected

During the establishment phase, pedigree information and performance records were collected on the Romney lambs born at Woodlands from 1970-1972. A maximum of 38 variables were recorded on each lamb from birth to 2-tooth stage including information on the sire and dam. The sire code denotes breed, source and year of use. Identification code, source, breed and age at the lamb's birth were recorded for the dam. Sex, date of birth and birth-rearing rank were recorded on the lamb. Date of birth was re-coded for the analysis from the beginning of the lambing season each year.

The liveweight and fleece data recorded on each lamb and the age at which the records were collected are shown in Table 3.2. The abbreviations for each trait used in subsequent chapters are also shown in the table.

Birthweight was recorded within 24 hours of birth. The lambs were actually docked at birth; the trait docking weight was recorded when the lamb was approximately 3-4 weeks old. However the age at measurement of this trait was relatively more variable than the other liveweights. Weaning weights were recorded on the same day for all lambs. After weaning, the 2 sexes were grazed separately and liveweights were recorded up to a week apart in the two groups. In a few of the later liveweights the 2 sexes were weighed up to 15 days apart.

Birthweight, docking weight and the greasy fleece weights were recorded to the nearest 0.01 kg., and the remaining liveweights were recorded to the nearest 0.1 kg.

Staple length was recorded to the nearest 0.1 cm from a mid-side sample collected at shearing.

Quality number was recorded as a mean value of the subjective score (e.g., 46/48 was coded as 47) assessed on the mid-side sample.

Character was recorded on a subjective scale from 1 (very bad) to 7 (excellent) based on an assessment of the overall fleece at shearing.

Greasy fleece colour was subjectively assessed at shearing on a scale from 1 (bad) to 5 (very good) over the whole fleece.

Break severity was graded on a subjective scale from 1 denoting a slight break on pulling the staple from each end, to 4 for a very bad break.

Break position was initially scored subjectively from 1 (near the staple butt) to 5 (near the staple tip). However, the assessment of break position was changed to a linear measurement from the staple butt during the establishment phase; this variable has been omitted from the analysis.

TABLE 3.2: AGE AT MEASUREMENT AND THE ABBREVIATIONS USED FOR THE LIVEWEIGHT AND FLEECE TRAITS

Trait	Abbreviations used	Mean age at measurement (days)		
		1970	1971	1972
Birthweight	BWT	-	-	-
Docking weight	DWT	33	25	35
Weaning weight	WWT	94	101	99
March/April weight	APR	165	173	169
May/June weight	JUN	235	248	235
July/August weight	AUG	286	285	281
October/November weight	NOV	400	416	397
Two-tooth weight	2TH	467	483	490
Lamb fleece weight	LFW	123	137	113
Hogget fleece weight	HFW	396	418	404
Staple length	STL	"	"	"
Quality number	QNO	"	"	"
Character	CHR	"	"	"
Colour	COL	"	"	"
Break severity	BRS	"	"	"

CHAPTER FOUR

STATISTICAL METHODS

## CHAPTER FOUR

STATISTICAL METHODSI. RESTRICTED MAXIMUM LIKELIHOOD ESTIMATION OF VARIANCE COMPONENTSA. The Model Used

Previous studies suggested the following effects should be considered for a model to describe the liveweight and fleece data recorded on lambs, namely; year of birth, sex, birth-rearing rank, age of dam, date of birth, dam and sire. All these effects, except dam, were included in the model for estimating the variance components.

The data analysed in this study are, in fact, part of a large-scale experiment as discussed in Chapter Three. Each year a random sample of dams was drawn from the total ewe population of the 3 locations involved in the experiment and re-allocated to Woodlands. Some of the lambs born in different years will therefore have the same dam. However, the possibility of correlated dam effects is not expected to be significant. Accordingly, dam effects were included in the residual variance and the residuals assumed to be randomly and independently distributed. In fact, to obtain REML estimates of the variance components the random effects in the model were assumed to be independently and identically, normally distributed.

The problem of determining which interactions between the fixed effects in the model to include, was overcome by using a sub-class means model. Thus the model to estimate the variance components for both the liveweight and the fleece data recorded on the Romney lambs at Woodlands was:

$$y_{ijklmn} = \mu_{iklm} + s_{ij} + bx_{ijklmn} + e_{ijklmn} \quad (4.1)$$

where,

$y_{ijklmn}$  is the observation on the  $n^{\text{th}}$  lamb born in the  $i^{\text{th}}$  year, with the  $j^{\text{th}}$  sire, the  $k^{\text{th}}$  sex, the  $l^{\text{th}}$  birth-rearing rank and of the  $m^{\text{th}}$  dam-age group.

$\mu_{iklm}$  is the mean of the observations in the  $iklm^{\text{th}}$  subclass.

$s_{ij}$  is the random effect of the  $j^{\text{th}}$  sire nested in the  $i^{\text{th}}$  year. Sire effects are assumed to be identically and independently, normally distributed with mean zero and variance  $\sigma_s^2$ .

$x_{ijklmn}$  is the (coded) date of birth of the  $ijklmn^{\text{th}}$  lamb.

$b$  is the coefficient of the regression of  $y_{ijklmn}$  on  $x_{ijklmn}$ .

$e_{ijklmn}$  is the random residual effect unique to the  $ijklmn^{\text{th}}$  observation. Residual effects are assumed to be identically and independently, normally distributed with mean zero and variance  $\sigma_e^2$ .

Furthermore, sire effects and residual effects are assumed to be uncorrelated.

Sires were fitted as a random effect, nested within years. Although there were 6 repeated sires, the data were not considered to be sufficiently connected for a cross-classified model. Accordingly, the records on the second set of progeny of these sires were deleted and a nested model adopted. The records on the progeny of a further 2 rams which were sons of sires used in the first year were also deleted; the assumption was then made that sire effects were random

and uncorrelated.

Date of birth was included as a covariate in the model. The overall regression was fitted for the variance components analysis, as fitting the within years regression was not expected to affect the estimation of the variance components to any great degree. Date of birth was re-coded from the beginning of the lambing season for each year so that date of birth reflected the actual age of the lamb at the time of recording.

The model may be expressed in matrix notation:

$$\underline{y} = \underline{X}\alpha + \underline{Z}b + \underline{e} \quad (4.2)$$

where

$\underline{y}$  is an  $N \times 1$  vector of observed records

$\underline{X}$  is a known  $N \times p$  design matrix

$\alpha$  is an unknown  $p \times 1$  vector of fixed effects; including the covariate, date of birth, and the subclass means of year, sex, birth-rearing rank and dam-age group.

$\underline{Z}$  is a known  $N \times q$  incidence matrix

$\underline{b}$  is a non-observable  $q \times 1$  vector of random sire effects which are assumed to be  $\sim N(0, \underline{I}\sigma_s^2)$

$\underline{e}$  is an unknown  $N \times 1$  vector of random residual effects which are assumed to be  $\sim N(0, \underline{I}\sigma_e^2)$

It is also assumed that,

$$\text{cov}(\underline{b}, \underline{e}) = \underline{0}$$

and that,

$$\underline{y} \text{ is } \sim N(\underline{X}\alpha, \underline{V}) \text{ where } \underline{V} = \underline{Z}\underline{Z}'\sigma_s^2 + \underline{I}\sigma_e^2.$$

### B. The Computing Algorithm Used

The REML estimates of the variance components were generated by iterating the equations given in (2.19) and (2.20). For the model of (4.2) with just one random effect, other than the residual, the equations which generate REML estimates are:

$$\hat{\sigma}_e^2(r+1) = \frac{\underline{y}'\underline{y} - \underline{y}'\underline{X}\hat{\underline{\alpha}}(r) - \underline{y}'\underline{Z}\hat{\underline{b}}(r)}{N - p^*} \quad (4.3)$$

and,

$$\hat{\sigma}_s^2(r+1) = \frac{(\hat{\underline{b}}'\hat{\underline{b}})(r)}{q - \text{tr}(\underline{T})} \quad (4.4)$$

where  $N$  is the total number of records on the trait being analysed,  $p^*$  is the rank of  $\underline{X}$ ,  $q$  is the number of sires,  $\underline{T}$  is the matrix defined in (2.21), which is estimated from the mixed model equations as shown in the appendix, and  $\hat{\underline{\alpha}}$  and  $\hat{\underline{b}}$  are solutions to the MME on each iteration. With  $\text{var}(\underline{b}) = \underline{I}\sigma_s^2$  and  $\text{var}(\underline{e}) = \underline{I}\sigma_e^2$ , Henderson's MME to be solved in this case are:

$$\begin{bmatrix} \underline{X}'\underline{X} & \underline{X}'\underline{Z} \\ \underline{Z}'\underline{X} & \underline{Z}'\underline{Z} + \underline{I}\sigma_e^2/\sigma_s^2 \end{bmatrix} \begin{bmatrix} \hat{\underline{\alpha}} \\ \hat{\underline{b}} \end{bmatrix} = \begin{bmatrix} \underline{X}'\underline{y} \\ \underline{Z}'\underline{y} \end{bmatrix} \quad (4.5)$$

Since solving (4.5) requires inversion of the left-hand-side matrix, the equations for the fixed effects were absorbed.

### C. Comparisons with the Method III and MIVQUE Estimates of the Variance Components

To compare REML estimates of the genetic and phenotypic parameters with estimates from other methods of estimation, the variance components were also estimated by Henderson's Method III and MIVQUE,

under the same model (4.2).

The MIVQUE's were obtained directly as the first iterate from the REML Fortran computing program.

The Method III estimates were obtained from REG, a generalised least squares program, which has been developed for use on the Prime 750 at Massey University by A.R. Gilmour.

## II. ESTIMATION OF HERITABILITY

Heritabilities were estimated from the paternal half-sib correlation using the REML estimates of the variance components. That is:

$$\hat{h}^2 = 4 \left( \frac{\hat{\sigma}_s^2}{\hat{\sigma}_s^2 + \hat{\sigma}_e^2} \right) \quad (4.6)$$

Blair (1981) gives a method for estimating the standard error of the heritability, SE ( $h^2$ ), that can be applied to any model, namely:

$$SE(h^2) = h^2 \sqrt{\frac{\text{var}(\sigma_s^2)}{(\sigma_s^2)^2} + \frac{\text{var}(\sigma_s^2) + \text{var}(\sigma_e^2) + 2\text{cov}(\sigma_s^2, \sigma_e^2)}{(\sigma_s^2)^2 + (\sigma_e^2)^2 + 2\sigma_s^2\sigma_e^2} - \frac{2\text{var}(\sigma_s^2) + \text{cov}(\sigma_s^2, \sigma_e^2)}{(\sigma_s^2)^2 + \sigma_s^2\sigma_e^2}} \quad (4.7)$$

The variances and covariances of the variance components in (4.9) are obtained from the variance-covariance matrix of  $\dot{\sigma}^2$ . That is:

$$\text{var}(\dot{\sigma}^2) = \text{var} \begin{bmatrix} \sigma_s^2 \\ \sigma_e^2 \end{bmatrix} = \begin{bmatrix} \text{var}(\sigma_s^2) & \text{cov}(\sigma_s^2, \sigma_e^2) \\ \text{cov}(\sigma_s^2, \sigma_e^2) & \text{var}(\sigma_e^2) \end{bmatrix} \quad (4.8)$$

Under REML,  $\hat{\text{var}}(\dot{\sigma}^2)$  is obtained from the following expression (see also Searle, 1970; Corbeil and Searle, 1976a):

$$\text{var}(\dot{\sigma}^2) = [\underline{I}(\dot{\sigma}^2)]^{-1} \quad (4.9)$$

where  $\underline{I}(\dot{\sigma}^2)$  is the information matrix derived by Searle (1979b).

For the model (4.2) this is:

$$\text{var}(\dot{\sigma}^2) = \frac{1}{2} \begin{bmatrix} \frac{N-p-q}{(\sigma_e^2)^2} + \frac{1}{\sigma_e^2} (\text{tr}[\underline{T}^2]) & \{(\text{tr}[\underline{T}] - \text{tr}[\underline{T}^2]) / \sigma_e^2 \sigma_s^2\} \\ \{(\text{tr}[\underline{T}] - \text{tr}[\underline{T}^2]) / \sigma_e^2 \sigma_s^2\} & \{(\text{tr}[(\underline{I}_q - \underline{T})^2]) / (\sigma_s^2)^2\} \end{bmatrix} \quad (4.10)$$

The elements of (4.10) correspond with those in (4.8). Accordingly estimates of these terms and the REML estimates of the variance components themselves, can be substituted in (4.7) to obtain  $\hat{SE}(\hat{h}^2)$ .

However, time and computing difficulties curtailed the estimation of these standard errors for the  $h^2$ 's. Using an approximation method instead, was not considered worthwhile.

Estimates of the Method III and MIVQUE variance component estimates can also be derived from (4.6), and compared to the REML estimates of  $h^2$ .

### III. ESTIMATION OF THE CORRELATIONS

The relationship between the paternal half-sibs is also used to estimate the correlations. The expression in (2.49) generates an estimate of the genetic correlation and (2.48) is used for estimating the phenotypic correlations. The REML estimates of the variance components are substituted into these expressions. However, instead of using sums of cross products, the covariance between two traits, 1 and 2, is obtained by making use of the well-known variance of a sum rule:

$$\hat{\sigma}_{1+2}^2 = \hat{\sigma}_1^2 + \hat{\sigma}_2^2 + 2 \hat{\sigma}_{1,2} \quad (4.11)$$

By setting up an additional variable (1 + 2) and after estimation of the variance components for all variables, (4.11) can then be rearranged to provide an estimate of the covariance. The REML estimates of the sire and residual variances of 1, 2 and (1 + 2) can be used to generate the sire and residual covariances. The phenotypic variance and covariance are then obtained as the sum of the sire and residual components.

Thus, the genetic correlation is estimated as:

$$\hat{r}_g = \frac{\frac{1}{2}(\hat{\sigma}_{s_{1+2}}^2 - \hat{\sigma}_{s_1}^2 - \hat{\sigma}_{s_2}^2)}{\sqrt{\hat{\sigma}_{s_1}^2 \cdot \hat{\sigma}_{s_2}^2}} \quad (4.12)$$

and the phenotypic correlation is estimated as:

$$\hat{r}_p = \frac{\frac{1}{2}(\hat{\sigma}_{s_{1+2}}^2 + \hat{\sigma}_{e_{1+2}}^2 - \hat{\sigma}_{s_1}^2 - \hat{\sigma}_{e_1}^2 - \hat{\sigma}_{s_2}^2 - \hat{\sigma}_{e_2}^2)}{\sqrt{(\hat{\sigma}_{s_1}^2 + \hat{\sigma}_{e_1}^2)(\hat{\sigma}_{s_2}^2 + \hat{\sigma}_{e_2}^2)}} \quad (4.13)$$

Standard errors were not estimated for either  $\hat{r}_g$  or  $\hat{r}_p$  because the extent to which approximations need to be made to account for lack of knowledge of the covariance terms, renders the standard error itself unreliable.

#### IV. ESTIMATION OF THE NON-GENETIC (FIXED) EFFECTS

##### A. The Models Used

Analyses of variance of the residuals of the liveweight data indicated that some heteroscedasticity was present. This was removed when the logarithms of the liveweight data were analysed. However, because the tests of significance of the main effects and interactions in the model were similar for both the liveweight and the log of the liveweight analyses, it was considered that the heteroscedasticity was unimportant for the purposes of estimating the environmental effects.

The lambs were maintained in one mob until weaning but separately thereafter. Therefore, birthweight, docking weight and weaning weight were analysed with both sexes together and the later liveweights and the fleece traits were analysed within-sex.

The covariate, date of birth, was found to be heterogeneous across years for the liveweights up to weaning. Accordingly, the regression of liveweight on date of birth was fitted within-years for birthweight, docking weight and weaning weight but the overall regression was fitted for the later liveweights and the fleece traits.

First-order interactions were only found to be significant in a few, isolated cases; hence all interactions were excluded from the models. Age of dam was found to have no effect on the fleece traits except lamb fleece weight and was therefore not included in the model for estimating environmental effects on the hogget fleece traits. Thus, 3 different models were fitted.

A general form of the mixed model is expressed in matrix notation in (4.2). The particular specifications of the model in this section

for estimating the fixed effects depend only on the composition of  $X\alpha$ . Thus,  $\alpha$ , the vector of unknown fixed effects, is now defined for the 3 models fitted.

To estimate the environmental effects for birthweight, docking weight and weaning weight,  $\alpha$  was defined as:

$$\alpha' = [\mu \quad \underline{t}' \quad \underline{f}' \quad \underline{r}' \quad \underline{d}' \quad \underline{b}'] \quad (4.14)$$

where,

$\mu$  is a general mean

$\underline{t}$  is a vector of year effects

$\underline{f}$  is a vector of sex effects

$\underline{r}$  is a vector of birth-rearing rank effects

$\underline{d}$  is a vector of dam-age group effects

$\underline{b}$  is a vector of within-year regression coefficients of the records on date of birth.

To estimate the environmental effects for lamb fleece weight and the later liveweights (from April through to 2-tooth weight) for each sex separately,  $\alpha$  was defined as:

$$\alpha' = [\mu \quad \underline{t}' \quad \underline{r}' \quad \underline{d}' \quad b] \quad (4.15)$$

where,

$b$  is the overall regression coefficient of the records on date of birth

and with,

the remaining terms in (4.15) as defined in (4.14).

To estimate the environmental effect for the hogget fleece traits, for each sex separately,  $\alpha$  was defined as:

$$\alpha' = [\mu \quad \underline{t}' \quad \underline{r}' \quad b] \quad (4.16)$$

with,

all the terms in (4.16) as previously defined.

When the vector of observations  $\underline{y}$  is assumed to be independently, normally distributed with mean  $\mu$  and variance  $\sigma^2$ , then the random variable:

$$U = \sum_{i=1}^N \frac{(y_i - \mu)^2}{\sigma^2} \quad (4.17)$$

has a chi-square distribution with N degrees of freedom. The F distribution is the distribution of the ratio of two independent chi-square variables divided by their respective degrees of freedom (Mood, Graybill and Boes, 1974).

Assuming the REML estimates of  $\sigma_s^2$  and  $\sigma_e^2$  to be the known parameter values, the main effects and interactions in the model were tested using an F-test of significance.

#### B. The Computing Algorithm Used

Assuming that the REML estimate of the vector of variance components,  $\hat{\underline{q}}^2$ , actually equals  $\underline{q}^2$ , the "obvious" REML estimator for  $\underline{q}$  given in (2.11) is identical to the best linear unbiased estimator (BLUE) for  $\underline{q}$ . The BLUE's of the fixed effects were estimated, using the computing package, REG, as:

$$\hat{\underline{q}} = (\underline{X}' \hat{\underline{V}}^{-1} \underline{X})^{-1} \underline{X}' \hat{\underline{V}}^{-1} \underline{y} \quad (4.18)$$

where,

$$\hat{\underline{V}}^{-1} = (\underline{Z}' \hat{\underline{Z}} \sigma_s^2 + \underline{I} \sigma_e^2)^{-1} \quad (4.19)$$

The standard error of the BLUE of  $\hat{\underline{q}}$  is estimated (see, for example,

Searle, 1971a) as:

$$\hat{SE}(\hat{\alpha}) = \sqrt{(\underline{X}'\hat{V}^{-1}\underline{X})^{-1}}$$

(4.20)

C H A P T E R F I V E

RESULTS

## CHAPTER FIVE

RESULTS

Estimates of the heritabilities for the liveweight and fleece traits are presented in Table 5.1. In many cases the estimates of  $h^2$  generated by the 3 methods are the same. The correspondence of the  $h^2$  estimates from the two methods is influenced to a large extent by the fact that the Method III and REML estimators of  $\sigma_e^2$  differ only in their denominators. The numerators of  $\hat{\sigma}_e^2$  are identical in both cases, namely, the total sum of squares (i.e.,  $\underline{y}'\underline{y}$ ) less the reduction in sums of squares due to fitting the full model, (i.e.,  $\underline{y}'\underline{X}\hat{\underline{\alpha}} + \underline{y}'\underline{Z}\hat{\underline{b}}$ ). REML uses  $N$  less the rank of  $\underline{X}$  in the denominator of  $\hat{\sigma}_e^2$ ; whereas Method III uses  $N$  less the rank of the full model (i.e.,  $r(\underline{X}) + r(\underline{Z})$ ). When  $N$  is large relative to  $(r(\underline{X}) + r(\underline{Z}))$ , the denominators will not be greatly different and therefore the REML and Method III estimates of  $\sigma_e^2$  will be very similar.

The Method III estimates were used as the a priori values for the MIVQUE estimates. Given the general similarity of the Method III and REML estimates, it is not surprising to find the MIVQUE estimates of  $h^2$  in Table 5.1 are also similar.

Standard errors were not estimated for the  $h^2$  estimates, as noted in Chapter Four. However, the  $h^2$  estimates from the 3 methods would be within the range of the standard errors since the  $h^2$  estimates themselves are the same in most cases.

The genetic and phenotypic correlations estimated with the REML estimates of the variance components are presented in Table 5.2. In all cases, except one (the estimate of the genetic correlation between

character and break severity), the estimates of the correlations fall within the parameter range from -1 to +1. The genetic correlation estimates are higher than the estimates of the phenotypic correlations in most cases; a trend commonly observed for the productive traits in sheep. The estimates also agree generally in terms of sign.

There are a few unexpectedly high estimates in the table (e.g., the estimate of 0.81 for the genetic correlation between character and staple length). It is noticeable that the few erratic correlation estimates that occur, involve traits that are assessed subjectively. Subjective assessment itself is well-known to be one of the causes of error variance and could inflate the correlation estimate.

The use of the F statistic for tests of significance of the fixed effects in the mixed models defined in (4.2) and (4.14) - (4.16) was discussed in Chapter Four. Tables 5.3 and 5.4 indicate that despite the presence of some heteroscedasticity, the significance of the main effects and interactions is practically identical for the liveweights and the logarithms of the liveweights. It was therefore considered justified to estimate fixed effects on the liveweight data.

First-order interactions were found to be significant in a few cases. The slope of the regression line was found to be significantly different for different dam-age classes on birthweight and for different birth-rearing rank classes on docking weight. However, they were non-significant in all other cases; hence, the only within-class regressions of date of birth fitted were the within-year regressions for birthweight, weaning weight and docking weight.

The interaction of birth-rearing rank x age of dam on birthweight was highly significant. An analysis of the effects indicated that

older ewes maintain the birthweight of singles relative to twins better than do younger ewes. Since the young (2-year-old) ewe is still growing, it is to be expected that she may be less able to "buffer" the birthweight of twins than an older ewe can.

The interaction of year x birth-rearing rank on weaning weight was highly significant. An analysis of the effects indicated that the weaning weights of twin lambs relative to singles were considerably lower in 1970 than the following 2 years. Given the adverse conditions that occurred in 1970 (see Chapter Three) compared to the following years, this interaction suggests that a poor year affects twins more severely than singles.

Since these interactions appear to be understandable and because the interactions were not systematic, they were excluded from the models for estimating the fixed effects. The BLUE's of the fixed effects under the different models (4.14), (4.15) and (4.16) are presented in Tables 5.6, 5.7 and 5.8. The means and standard deviations for each trait and the numbers of records analysed are also listed in the tables. The term not applicable (na) appears in the table whenever a non-genetic effect was not fitted for a particular trait.

**TABLE 5.1:** ESTIMATES OF THE HERITABILITIES FROM DIFFERENT METHODS OF VARIANCE COMPONENT ESTIMATION

Trait <sup>1</sup>	No. of records	Heritability estimates		
		REML	Method III	MIVQUE <sup>2</sup>
BWT	1602	0.08	0.08	0.09
DWT	1598	0.05	0.05	0.05
WWT	1559	0.10	0.10	0.10
APR	1511	0.13	0.13	0.13
JUN	1455	0.14	0.14	0.14
AUG	1427	0.15	0.15	0.15
NOV	1398	0.24	0.25	0.24
2TH	1377	0.31	0.32	0.32
LFW	1560	0.19	0.19	0.19
HFW	1397	0.30	0.30	0.30
STL	1397	0.37	0.38	0.37
QNO	1396	0.56	0.55	0.56
CHR	1387	0.19	0.19	0.18
COL	1387	0.26	0.26	0.25
BRS	1209	0.07	0.08	0.09

1. Abbreviations given in Table 3.2.

2. Using the Method III estimates of  $\sigma_e^2$  and  $\sigma_s^2$  as a priori values.

TABLE 5.2: ESTIMATES OF THE GENETIC (ABOVE DIAGONAL) AND PHENOTYPIC (BELOW DIAGONAL) CORRELATIONS BETWEEN THE LIVELWEIGHT AND FLEECE TRAITS

Trait	BWT	DWT	WWT	APR	JUN	AUG	NOV	2TH	LFW	HFW	STL	QNO	CHR	COL	BRS
BWT		0.79	0.78	0.25	0.04	0.02	0.31	0.18	0.80	0.56	0.07	-0.10	0.27	0.23	-0.43
DWT	0.47		0.50	0.22	-0.01	0.25	0.21	0.32	0.47	0.40	-0.40	0.14	-0.42	-0.27	-0.01
WWT	0.42	0.64		0.80	0.38	0.49	0.45	0.50	0.81	0.48	0.02	0.04	0.00	-0.05	-0.21
APR	0.36	0.53	0.80		0.86	0.91	0.85	0.93	0.62	0.78	0.16	0.14	-0.06	-0.50	-0.27
JUN	0.31	0.47	0.68	0.83		0.97	0.84	0.85	0.26	0.70	0.18	0.16	0.06	-0.38	-0.16
AUG	0.35	0.50	0.71	0.85	0.86		0.84	0.91	0.28	0.60	0.07	-0.01	-0.07	-0.39	0.15
NOV	0.29	0.39	0.57	0.68	0.69	0.74		0.91	0.32	0.54	0.45	0.14	0.30	-0.10	0.02
2TH	0.27	0.38	0.57	0.68	0.69	0.74	0.84		0.09	0.50	0.19	0.24	0.06	-0.34	0.22
LFW	0.36	0.50	0.65	0.52	0.46	0.48	0.35	0.32		0.59	0.35	-0.31	0.38	0.01	-0.68
HFW	0.13	0.14	0.26	0.42	0.44	0.48	0.47	0.42	0.32		0.74	-0.33	0.64	-0.03	-0.90
STL	-0.05	-0.08	-0.02	0.12	0.17	0.15	0.18	0.16	0.08	0.37		-0.43	0.81	0.09	-0.38
QNO	0.07	0.12	0.08	0.01	0.02	0.04	0.02	0.05	-0.05	-0.16	-0.34		-0.40	-0.13	0.70
CHR	0.03	0.07	0.08	0.17	0.23	0.24	0.25	0.18	0.15	0.35	0.32	-0.02		0.44	-1.01
COL	0.05	0.08	0.06	0.09	0.10	0.08	0.11	0.08	0.09	0.17	0.11	0.02	0.37		-0.22
BRS	0.03	0.03	-0.01	0.05	-0.11	-0.12	-0.08	-0.06	-0.05	-0.22	-0.17	0.14	-0.18	-0.08	

**TABLE 5.3: SIGNIFICANCE OF THE ENVIRONMENTAL EFFECTS ON THE LIVeweIGHTS**

1	BWT	DWT	WWT	APR		JUN		AUG		NOV		2TH	
				F	M	F	M	F	M	F	M	F	M
YR	**	**	**	**	**	**	**	**	**	**	**	**	**
SEX	**	**	**	na	na								
BRR	**	**	**	**	**	**	**	**	**	*	**	**	NS
AOD	**	**	**	**	**	**	**	**	**	**	NS	**	NS
w <sub>r</sub> <sub>DOB</sub>	**	**	**	na	na								
r <sub>DOB</sub>	na	na	na	**	**	**	**	**	**	**	**	**	**
DOBxAOD	**	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
BRRxAOD	**	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
BRRxDOB	NS	**	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
YRxBRR	NS	NS	**	NS	NS								

1. Abbreviations given in Table 5.7 for the main effects; following these are the interaction terms that were significant for any liveweight.

\*\* Significant at the 1% level.

\* Significant at the 5% level.

NS non-significant.

na not applicable - see text.

TABLE 5.4: SIGNIFICANCE OF THE ENVIRONMENTAL EFFECTS ON THE LOGARITHM OF THE LIVEWEIGHTS

1	L(BWT)	L(DWT)	L(WWT)	L (APR)		L (JUN)		L (AUG)		L (NOV)		L (2TH)	
				F	M	F	M	F	M	F	M	F	M
YR	**	**	**	**	**	**	**	**	**	**	**	**	*
SEX	**	**	**	na	na								
BRR	**	**	**	**	**	**	**	**	**	**	**	**	NS
AOD	**	**	**	**	**	**	**	**	**	**	NS	**	NS
w <sub>r</sub> <sub>DOB</sub>	**	**	**	na	na								
r <sub>DOB</sub>	na	na	na	**	**	**	**	**	**	**	**	**	*
DOB x AOD	*	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
BRR x AOD	*	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
YR x BRR	NS	NS	**	NS	NS								
BRR x DOB	NS	*	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS

1. Abbreviations given in Table 5.7 for the main effects; following these are the interaction terms that were significant for any trait.

\*\* Significant at the 1% level.

\* Significant at the 5% level.

NS non-significant.

na not applicable - see text.

**TABLE 5.5: SIGNIFICANCE OF THE ENVIRONMENTAL EFFECTS ON THE FLEECE TRAITS**

1	LFW		HFW		STL		QNO		CHR		COL		BRS	
	F	M	F	M	F	M	F	M	F	M	F	M	F	M
YR	**	**	**	**	**	**	**	**	**	NS	**	**	**	**
BRR	**	**	NS	*	**	*	NS	NS	NS	NS	NS	NS	NS	NS
AOD	**	**	NS	NS										
r <sub>DOB</sub>	**	**	**	**	**	**	**	**	**	*	NS	NS	**	**
YRxBRR	NS	**	NS	NS										

1. Abbreviations given in Table 5.7 for the main effects; following these is the only interaction to be significant for any fleece trait.

\*\* Significant at the 1% level.

\* Significant at the 5% level.

NS non-significant.

TABLE 5.6: ESTIMATES OF THE ENVIRONMENTAL EFFECTS (KG) ON THE LIVEWEIGHTS WITH BOTH SEXES COMBINED

	Abbrevi- ation	BWT	DWT	WWT
Mean (kg)	$\bar{x}$	4.39	10.08	21.08
Standard deviation (kg)	SD	1.09	2.76	4.24
No. of records	N	1537	1534	1495
Year of Birth	YR			
1972-1970	72-70	-0.25 ± 0.11	-0.45 ± 0.25	-0.81 ± 0.43
1972-1971	72-71	-0.65 ± 0.12	4.54 ± 0.26	-4.28 ± 0.45
Sex	SEX			
ram-ewe	M-F	0.28 ± 0.05	0.68 ± 0.10	+1.68 ± 0.17
Age of Dam	AOD			
5-2 year olds	5-2	0.54 ± 0.07	1.11 ± 0.16	1.68 ± 0.26
5-3 year olds	5-3	-0.02 ± 0.06	-0.26 ± 0.14	-0.26 ± 0.24
5-4 year olds	5-4	-0.12 ± 0.07	-0.38 ± 0.15	-0.45 ± 0.24
Birth/rearing rank	BRR			
single/single -				
twin/single	SS-TS	0.93 ± 0.09	1.42 ± 0.19	1.45 ± 0.32
single/single -				
twin/twin	SS-TT	0.93 ± 0.05	2.65 ± 0.16	3.59 ± 0.19
Within-year re- gression of weight on date of birth	$w_{r_{DOB}}$			
1970	70	0.01 ± 0.01	-0.13 ± 0.01	-0.21 ± 0.02
1971	71	0.04 ± 0.01	-0.25 ± 0.01	-0.12 ± 0.02
1972	72	0.03 ± 0.01	-0.13 ± 0.01	-0.16 ± 0.02

TABLE 5.7: ESTIMATES OF THE ENVIRONMENTAL EFFECTS (KG) ON THE LIVeweIGHTS FOR EACH SEX SEPARATELY

1	APR		JUN		AUG	
	F	M	F	M	F	M
$\bar{x}$	22.82	26.66	24.77	27.52	26.39	25.52
SD	4.05	5.04	3.95	5.25	4.03	4.53
N	710	672	709	617	708	590
YR						
72-70	-1.70 ± 0.32	3.63 ± 0.42	1.78 ± 0.32	7.65 ± 0.44	4.11 ± 0.33	6.53 ± 0.76
72-71	-5.25 ± 0.34	-1.65 ± 0.44	-3.42 ± 0.35	2.77 ± 0.45	-1.86 ± 0.35	1.12 ± 0.75
BRR						
SS-TS	1.10 ± 0.46	0.65 ± 0.54	0.80 ± 0.45	0.30 ± 0.55	0.90 ± 0.46	0.43 ± 0.51
SS-TT	2.24 ± 0.26	3.09 ± 0.34	1.75 ± 0.26	2.27 ± 0.35	1.74 ± 0.26	2.18 ± 0.32
AOD						
5-2	0.47 ± 0.36	1.31 ± 0.46	0.07 ± 0.36	0.99 ± 0.48	0.25 ± 0.35	0.72 ± 0.44
5-3	-0.62 ± 0.32	-0.71 ± 0.43	-0.91 ± 0.32	-0.42 ± 0.44	-0.90 ± 0.32	-0.44 ± 0.40
5-4	-0.49 ± 0.33	-1.02 ± 0.43	-0.67 ± 0.33	-0.67 ± 0.44	-0.88 ± 0.33	-0.74 ± 0.40
$r_{DOB}^2$	-0.13 ± 0.01	-0.19 ± 0.02	-0.13 ± 0.01	-0.18 ± 0.02	-0.12 ± 0.01	-0.20 ± 0.03

1. Abbreviations given in Table 5.6.

2. The overall regression (across years) of weight on date of birth was fitted for the later liveweights.

TABLE 5.7: (continued)

1	NOV		2TH	
	F	M	F	M
$\bar{x}$	30.68	32.62	41.98	43.86
SD	4.23	5.05	5.29	6.53
N	699	570	686	562
YR				
72-70	1.32 ± 0.44	-4.56 ± 0.94	6.08 ± 0.59	-3.84 ± 0.68
72-71	-3.58 ± 0.46	4.26 ± 0.92	0.30 ± 0.62	8.64 ± 0.69
BRR				
SS-TS	0.52 ± 0.57	0.03 ± 0.61	0.32 ± 0.71	0.12 ± 0.76
SS-TT	1.05 ± 0.31	1.62 ± 0.39	1.38 ± 0.40	1.03 ± 0.47
AOD				
5-2	0.32 ± 0.43	0.92 ± 0.53	0.31 ± 0.54	1.42 ± 0.65
5-3	-0.69 ± 0.39	0.05 ± 0.48	1.56 ± 0.49	0.25 ± 0.59
5-4	1.01 ± 0.40	-0.38 ± 0.49	1.30 ± 0.50	0.00 ± 0.59
$r_{DOB}^2$	-0.10 ± 0.01	-1.11 ± 0.03	-0.10 ± 0.02	-0.12 ± 0.03

1. Abbreviations given in Table 5.6.

2. The overall regression (across years) of weight on date of birth was fitted for the later liveweights.

TABLE 5.8: ESTIMATES OF THE ENVIRONMENTAL EFFECTS (KG) ON THE FLEECE TRAITS FOR EACH SEX SEPARATELY

1	LFW		HFW		STL	
	F	M	F	M	F	M
$\bar{x}$	1.21	1.23	1.86	1.84	10.99	10.05
SD	0.38	0.40	0.54	0.52	1.89	1.61
N	730	701	697	571	697	571
YR						
72-70	-0.10 ± 0.03	-0.14 ± 0.03	-0.74 ± 0.04	-0.80 ± 0.04	2.13 ± 0.22	1.03 ± 0.22
72-71	-0.55 ± 0.03	-0.68 ± 0.03	-1.52 ± 0.04	-1.22 ± 0.05	-0.49 ± 0.23	0.08 ± 0.22
BRR						
SS-TS	0.14 ± 0.04	0.10 ± 0.03	0.06 ± 0.04	0.04 ± 0.05	-0.25 ± 0.25	0.04 ± 0.23
SS-TT	0.26 ± 0.02	0.29 ± 0.02	0.02 ± 0.02	0.09 ± 0.03	-0.48 ± 0.14	-0.34 ± 0.14
$r_{DOB}^2$	-0.10 ± 0.001	-0.01 ± 0.001	-0.01 ± 0.001	-0.01 ± 0.002	-0.03 ± 0.01	-0.02 ± 0.008
AOD <sup>3</sup>						
5-2	0.10 ± 0.03	0.13 ± 0.03	na	na	na	na
5-3	-0.02 ± 0.03	-0.03 ± 0.03	na	na	na	na
5-4	-0.03 ± 0.03	-0.04 ± 0.03	na	na	na	na

1. Abbreviations given in Table 5.6.

2. Overall regression coefficient of the fleece traits on date of birth.

3. Age of dam not fitted in model for the hogget fleece traits.

na. not applicable - see text.

TABLE 5.8: (continued)

1	QNO		CHR		COL		BRS	
	F	M	F	M	F	M	F	M
$\bar{x}$	48.02	47.65	4.01	4.14	4.03	4.13	2.18	3.11
SD	1.90	2.01	0.80	0.77	0.66	0.65	0.99	1.33
N	698	569	690	568	695	563	547	533
YR								
72-70	-1.63±0.28	-0.17±0.31	-1.08±0.08	0.19±0.09	-0.89 ±0.07	-0.51±0.08	0.87±0.10	1.99±0.11
72-71	-0.71±0.29	-1.61±0.32	-1.04±0.08	0.16±0.09	-0.85 ±0.07	-0.02±0.08	1.02±0.09	1.60±0.11
BRR								
SS-TS	-0.46±0.28	-0.10±0.28	-0.01±0.10	0.08±0.11	0.10 ±0.09	0.05±0.09	-0.02±0.15	0.18±0.15
SS-TT	0.09±0.15	0.26±0.18	-0.05±0.05	0.09±0.07	0.01 ±0.05	0.07±0.06	0.11±0.08	0.09±0.09
$r_{DOB}^2$	-0.04±0.01	0.03±0.01	-0.01±0.003	-0.01±0.004	-0.004±0.002	-0.01±0.003	0.01±0.003	0.01±0.01
AOD <sup>3</sup>								
5-2	na	na	na	na	na	na	na	na
5-3	na	na	na	na	na	na	na	na
5-4	na	na	na	na	na	na	na	na

1. Abbreviations given in Table 5.6.

2. Overall regression coefficient of the fleece traits on date of birth.

3. Age of dam not fitted in model for the hogget fleece traits.

na. not applicable - see text.

CHAPTER SIX

DISCUSSION

## CHAPTER SIX

DISCUSSIONI. NON-GENETIC EFFECTS

The estimation of non-genetic effects is useful for explaining some of the variation between the phenotypic records of individuals and also to adjust records for these effects prior to their use in estimating B.V.'s. Clarke and Rae (1976) observe that the increase in accuracy of the B.V. estimate that may be gained from adjusting for these non-genetic effects, depends on the accuracy with which the effects themselves can be estimated. This, in turn, depends on how well the different environmental levels can be identified and the animals divided into these classes (e.g., twins and singles; rams and ewes), how well the effect can, in fact, be estimated for each of these classes and how much the effect actually varies between individuals within each class.

In this study the non-genetic effects estimated from each of the 3 models defined in Chapter Four, were found to be significant for most of the traits described by the model. The best linear unbiased estimates (BLUE) of the environmental effects are presented in Tables 5.6, 5.7 and 5.8 for the 3 models used. The standard errors are relatively high in many cases, particularly in the few cases where a non-significant effect has been estimated (e.g., the effect of SS-TS on 2-tooth liveweight is  $0.12 \pm 0.76$  for the rams, although in this specific case BRR is shown to be non-significant in Table (5.3)).

Year-of-birth effects, although significant, are specific to the flock and year of study in which they are measured, and cannot be

generalised (as correction factors in Sheeplan, for example). Nevertheless, it is interesting to note that, although 1970 had a large adverse effect on the lambs growth relative to 1971, particularly in the early stages of the lambs' growth, that 1972, in fact, had an even greater negative influence, although the difference is much smaller. This same trend is generally evident for the fleece weights in Table (5.8). The fleece quality traits also seem to have been most adversely affected in 1972 and least in 1971. There is also a noticeably larger effect of year of birth on the ram than the ewe lambs for the fleece traits and the later liveweights.

The effect of sex on birthweight, docking weight and weaning weight, when the lambs were still being grazed together, is shown in Table 5.6. As expected, the ram lambs have an advantage for liveweight at these ages. The REML estimates obtained here are somewhat lower than previous estimates reviewed in Table 2.2 for weaning weight. There do not appear to be any previous estimates of environmental effects on birthweight or docking weight.

From Tables 5.6 and 5.7 it appears that until August (10 month) liveweight lambs with a young dam are at a disadvantage relative to lambs with older dams. After August, the effect of age of dam is not as large nor as consistent. It is also apparent that a 5-year-old ewe is a disadvantage to the lamb relative to 3 and 4 year-old dams.

By subtracting the estimates for the (5-4) year-old ewe from the (5-3) and (5-2) year-old age-of-dam effects in Tables 5.6 - 5.8 an indication of dam-age effects relative to the 4 year-old ewe may be obtained. These can then be compared with the age-of-dam effects from previous studies (in Table 2.2) and those used for Sheeplan (in

Table 2.1), since the mature ewe class largely comprises 4 year-old ewes.

There seems to be general agreement of the BLUE's of age-of-dam effects on weaning weight from this study and the estimates previously obtained. The "derived" BLUE's of 1.23 for (4-2) year-old ewes and 0.19 for (4-3) year-old dam-age effects, agree well with the values of 1.3 and 0.2, respectively, which are currently used by Sheeplan.

The estimates of environmental effects after weaning have been made within-sex in this study, although this is not the case for the other estimates reviewed in Tables 2.1 and 2.2. It appears that the corresponding BLUE's of age-of-dam effects on liveweight are lower for the females, and higher for the males, than the previous combined-sex estimates. However, the standard errors are quite high relative to the BLUE's. Sheeplan currently makes the same adjustment to hogget liveweight for age-of-dam effects to both sex classes in dual-purpose breeds.

The adjustments made to the post-hogget liveweights for age-of-dam effects are smaller, in relation to the size of the animals. This reflects the lessening influence of the dam with increasing age of the progeny. The BLUE's obtained in this study for 2-tooth weight indicate age of dam has a relatively small effect by this stage. In fact, Table 5.3 indicates these effects are non-significant for ram lambs both as hoggets and as 2-tooths.

Hogget fleece traits are not significantly influenced by age-of-dam effects, as Table 5.5 shows, and these were not, therefore, included in the model. Sheeplan also makes no adjustment to hogget fleece weight for age-of-dam effects. Lamb fleece weight is, however, significantly

influenced by age-of-dam effects. The BLUE's are slightly larger for males than females. There does not appear to be any previous estimates to compare with in the Romney.

Birth-rearing rank has a significant effect on the liveweights and fleece weights up to the hogget stage. The effects on 2-tooth weight and the hogget fleece quality traits are generally less significant.

The BLUE's of the effects of birth-rearing rank on weaning weight agree well with the findings from previous studies. The BLUE's for hogget liveweight for each sex separately are generally lower in both cases than the previous estimates obtained with the 2 sexes combined. The effect of birth-rearing rank on ewe hoggets estimated in this study appears to be particularly large and is, in fact, higher than the estimates in Table 2.2.

The effect of birth-rearing rank on liveweight also decreases with increasing age. Sheeplan makes the same adjustment to liveweight records for birth-rearing rank effects to both sexes. The BLUE's obtained here for April and August liveweight are generally slightly lower for both sexes than the estimates Sheeplan uses for the corresponding autumn and winter liveweights. The effect of birth-rearing rank declines, relative to the weight of the animal, with increasing age. In the males, birth-rearing rank was not found to have a significant effect on 2-tooth weight in this study.

Birth-rearing rank has a significant effect on lamb fleece weight with the effect being similar for both sexes. The effect on the hogget fleece traits is more variable and frequently non-significant as Table 5.5 shows. The BLUE's for hogget fleece weight in Table 5.8 are lower than the estimates used by Sheeplan, however, they are in

fairly close agreement with the estimates of Baker et al. (1974). The difference between the sexes for birth-rearing rank effects on the fleece traits is quite variable; in some cases the effect is greater on the ram and less on the ewe, in other cases it is greater for the ram, and in a number of cases it is negative in one sex and positive in the other. The estimates of birth-rearing rank effects on fleece quality traits obtained by Chopra (1978), although not reported here, also tended to be variable.

The BLUE's for the regression of the liveweight and fleece traits on date of birth of the lamb were found to be significant in most cases (Tables 5.3 and 5.5). In comparing these estimates with the results from the previous studies reported in Table 2.2 and from Sheeplan in Table 2.1, note that the regressions in these tables are expressed as the regression of liveweight on the age of the lamb. Since date of birth is equivalent to negative age, the estimates in Tables 2.1 and 2.2 will be opposite in sign to the BLUE's in Tables 5.6 - 5.8.

The within-year regression coefficients were fitted for birthweight through to weaning weight. The estimates pertaining to weaning weight generally agree with the estimates in Tables 2.1 and 2.2, although the 1970 regression is higher than the range of previous estimates. There is also a significant, and relatively large, difference between the within-year regression coefficients from the 3 years for weaning weight in Table 5.6. This is also evident amongst the within-year coefficients for birthweight and docking weight. Note that the regression of birthweight on date of birth is opposite in sign to the regressions for all the other liveweights. This is possibly caused by ewes carrying lambs later in the season, when feed supplies are better; thus, being able to

divert extra resources to the foetus in the late stages of gestation. Although gestation length is fairly stable in the ewe, there may also be some increase in birthweight associated with a longer gestation period.

The overall regression coefficients on date of birth were obtained for the later liveweights and fleece traits on a within-sex basis. There is generally good agreement between the BLUE's for the regression of hogget fleece weight on date of birth with the Sheeplan estimates in Table 2.1 and the estimates reported from previous studies in Table 2.2.

There is, however, some disagreement between the estimates in these tables for the regression of the later liveweights on date of birth and the corresponding BLUE's obtained in this study. There are large within-sex differences in the BLUE's which are not evident in the within-sex estimates used by Sheeplan, nor in the combined-sex estimates in Table 2.2. The BLUE for the regression of hogget fleece weight on date of birth in ram lambs is particularly different. Further evidence of such differences would be required to substantiate them.

In general, the BLUE's of this study have agreed reasonably well with previous estimates.

## II. HERITABILITY

The heritability estimates for the traits under study are presented in Table 5.1. As discussed in Chapter Five, the 3 methods (REML, Henderson's Method III and MIVQUE) used to estimate  $h^2$  are in good agreement; it is the values of the estimates themselves that are of interest in this chapter.

Growth traits in the Romney are considered to have medium to high heritabilities (Dalton and Rae, 1978). The estimates of  $h^2$  for the liveweight traits that are used in Sheeplan are based on the results of some of the earlier studies in Table 2.3 (e.g., Ch'ang and Rae, 1961, 1970). The  $h^2$  estimates of the liveweight traits in Table 5.1 from the present study are lower than many of these previous estimates. It is, however, a noticeable trend in Table 2.3 that in more recent years the  $h^2$  estimates of liveweight traits in the New Zealand Romney (e.g., Baker et al., 1974; Eikje, 1978b; Baker et al., 1979) are also lower and in many cases are in good agreement with the estimates of this study. The  $h^2$  estimates of liveweight in the Polish Romney given in Table 2.3 (e.g., Radomska and Klewiec, 1975, 1976) are also generally lower than the early New Zealand Romney estimates.

There have not been many estimates for  $h^2$  of birthweight in the Romney. The REML estimate of 0.08 is considerably lower than the range of estimates (from 0.19-0.35) from Blair (1981) for the New Zealand Romney, but is similar to the estimate of 0.013 from Radomska and Klewiec (1975) for the Polish Romney. Further estimates are required before a trend can be substantiated although it appears likely that birthweight has a low-medium  $h^2$ . Dalton and Rae (1978) suggest it would be useful to have more knowledge of the  $h^2$  of birthweight.

Several studies (e.g., Hight and Jury, 1970; McCutcheon, 1981) have shown that birthweight has an important effect on lamb mortality rates and lambing difficulty.

The particularly low estimate of  $h^2$  for docking weight of 0.05 in Table 5.1 is probably largely due to the wide variation in age at docking discussed in Chapter Three. This would increase  $\hat{\sigma}_e^2$  relative to  $\hat{\sigma}_s^2$  and therefore decrease  $\hat{h}^2$ . There appear to be no previous estimates of  $h^2$  at docking. Variability of environmental factors, such as the age of the lamb, are typical at docking in New Zealand; thus docking weight is not regarded as a useful trait for predicting B.V., nor as a basis for selection.

In fact, both birthweight and docking weight are recorded under variable conditions in New Zealand. A more uniform environment (e.g., if animals are housed at lambing; as is done in some other countries) would be expected to increase  $\hat{h}^2$ .

Until after weaning, the young lamb is particularly influenced by the maternal environment. Thus, DOR (including DDR) estimates of  $h^2$  can be expected to be higher than PHS estimates; as discussed in Chapter Two (section IIB). The DOR  $h^2$  estimates of weaning weight are all about 0.2 while most of the PHS estimates fall in the 0.1-0.2 range, including the REML estimate of 0.1 from this study. Eikje (1978b) observes that the PHS estimate is a better measure of the  $h^2$  of the lamb's own genetic potential and that individual and maternal effects on weaning weight should be considered as 2 separate traits.

Sheeplan also uses information, if available, on autumn liveweight (March-August) and winter liveweight (May-August) for predicting B.V. for weaning weight, although these are of less importance than weaning

weight itself and spring liveweight (September–November). The REML estimates of  $h^2$  in Table 5.1 for liveweight from April to August, when the lambs are, on average, 7 through to 10 months old, are generally lower than the corresponding estimates in Table 2.3 from other studies. The trend for  $h^2$  to increase as the lamb gets older is maintained, with the REML estimates increasing from 0.13–0.15 during this period.

Hogget liveweight, or spring liveweight as it is called in Sheeplan (Clarke and Rae, 1976), is another important trait in dual-purpose breeds such as the Romney. Hogget liveweight is regarded as having a high  $h^2$ ; Sheeplan uses a value of 0.35. The REML estimate for the  $h^2$  of hogget liveweight obtained in this study of 0.24 is noticeably lower. It does, however, agree well with some of the later estimates in Table 2.3 (e.g., Baker *et al.*, 1974; Baker, 1977; Baker *et al.*, 1979). This appears to be a similar trend as was observed with weaning weight. Given the particular importance of these 2 liveweights as selection criteria in the Romney, it is necessary to have reliable estimates of  $h^2$  available for use in selection plans.

The estimate for  $h^2$  of 2-tooth liveweight in Table 5.1 is also somewhat lower than would be expected from previous studies. Rae (pers. comm.) comments that estimates have usually been in the 0.4–0.5 range in contrast to the REML estimate of 0.31 obtained in this study. Although few estimates have actually been reported for 2-tooth body weight itself, Blair (1981) did find that the  $h^2$  of liveweight based on average lifetime records of the ewe were in the range of 0.26–0.48.

It is evident from this discussion of the  $h^2$  estimates for liveweight that, although they tend to be 0.1–0.2 units lower than many of the estimates in Table 2.3, they are generally in closer agreement with

some of the more recent estimates. Furthermore, the trend of increasing  $h^2$  of liveweight with increasing age of the lamb is clearly supported by this study.

The estimates of  $h^2$  amongst the fleece traits tend to be higher with the measurable traits, such as the fleece weights, and lower for the subjectively assessed traits such as character. Colour is an exception; having a moderate  $h^2$  estimate despite being an overall subjective assessment for the colour of the fleece as a whole.

There do not appear to be any previous estimates of  $h^2$  of lamb fleece weight available in the literature. It is therefore impossible to compare the estimate obtained in this study using REML estimates of the variance components. The estimate of 0.19 represents a fairly low heritability.

Hogget fleece weight is a particularly important trait in the Romney and it is therefore an advantage that the  $h^2$  estimate obtained here of 0.3, has been substantiated by many of the previous studies in Table 2.3. These estimates are in fact relatively stable compared to the  $h^2$  estimates for the liveweights and other fleece traits.

In the Romney, traits associated with fleece quality are less important than fleece weight, however, certain quality traits do have some influence on price. McPherson (1982) found that the effect of style grade on price was governed almost exclusively by the colour of the wool despite the fact that style is purported to reflect character, medullation, unevenness and other factors as well (Wickham, 1982a). Whiteness of the wool is particularly important in some of the end-uses of Romney wool such as carpet manufacture. The estimate of  $h^2$  of colour of 0.26 is comparable with most of the estimates in Table 2.3

which suggests colour may be reasonably heritable.

The estimate of 0.56 for the  $h^2$  of quality number also compares with other estimates for this trait in Table 2.3. The subjective assessment of quality number as a measure of wool fineness is now being discarded in preference to objective assessment of fibre diameter. Although the relationship between these 2 traits is not direct, the genetic and phenotypic correlations are commonly in the medium to high range (ross, 1964; Sumner, 1969; Blair, 1981); although some low estimates of the correlation have also been obtained (e.g., Chopra, 1978). Thus, quality number is often used as an indicator of fibre diameter.

Character and break severity both have low  $h^2$  estimates in Table 5.1. The estimates of  $h^2$  for character from previous studies in Table 2.3 vary markedly. However, neither of these two traits are important selection criteria for the Romney.

### III. CORRELATIONS

Correlations between different traits are particularly important in a dual-purpose breed, such as the Romney, where improvement in several productive characteristics are required simultaneously. In the Romney, both high fleece weight and increased meat production are important objectives. High levels of meat production are dependent on high growth rates and increased reproductive performance. In this study, liveweight and fleece data have been analysed. The genetic correlations between the liveweight traits and fleece weight are of particular interest in the Romney.

Improvement through selection is also enhanced by decreasing the generation interval. Wickham (1982b) stresses the importance of early breeding in sire-producing flocks. In this respect, the phenotypic correlations between the liveweight traits and between the fleece weight traits are important. Early selection of individuals for say, lamb fleece weight, will be of little advantage if the correlations (both phenotypic and genetic) between lamb fleece weight and hogget fleece weight are low.

The genetic and phenotypic correlations estimated using REML for the liveweight and fleece traits are presented in Table 5.3<sup>2</sup>. The phenotypic correlations between the liveweights are generally high or medium-high. The genetic correlations between the liveweights are consistently high from weaning, although those involving birthweight and docking weight are more variable.

There are few estimates of either genetic or phenotypic correlations between birthweight or docking weight and other traits, and appear to be none for the New Zealand Romney. Radomska and Klewiec (1975) have

estimated correlations of birthweight with weaning weight, hogget liveweight and greasy fleece weight. Their estimates of the phenotypic correlations are in all cases lower than the estimates from this study. It is difficult, however, to compare with their genetic correlation estimates involving birthweight since these are outside the -1 to +1 range for a correlation in 2 of the 3 cases. If the correlation between birthweight and weaning weight is as high as this study suggests then more use could be made of birthweight as an early selection criterion in situations where weighing lambs at birth is practical.

The estimates of the genetic correlations between weaning weight and later weights in Table 5.2 are generally lower than the previous estimates reviewed in Table 2.5, although still in the medium-high range. Given the importance of weaning weight as a selection objective this is an important advantage. The genetic correlations of weaning weight with lamb and hogget fleece weight are also strong and positive; another very useful trend in terms of improvement. The estimates of the phenotypic correlations between weaning weight and the later liveweights and fleece weights are also high and agree well with the estimates in Table 2.4.

Hogget liveweight is an important liveweight trait that is incorporated in the selection index of Sheeplan. It is therefore useful to note that hogget liveweight is not only genetically and phenotypically highly correlated with 2-tooth weight, but also with hogget fleece weight. Weaning weight, hogget fleece weight and number of lambs born are the other 3 traits in the index. The estimates of the genetic and phenotypic correlations between hogget liveweight and the other liveweights in Table 5.2 are generally comparable with the previous estimates

in Table 2.4.

The estimates of the genetic correlations between staple length and the liveweights in Table 5.2 is somewhat variable although always positive. The corresponding estimates in Table 2.5 tend to be positive and medium to strong. The phenotypic correlation estimates between staple length and liveweight are low and variable, which is also found for the estimates presented in Table 2.4.

The other fleece traits studied, are subjectively assessed traits and are of less importance as breeding objectives. Quality number is an exception insofar as it reflects fibre diameter, which is relatively important in the Romney. Quality number, character and break severity have low and variable estimates of their phenotypic and genetic correlations with the liveweights. The phenotypic correlation estimates involving colour are also low and variable, however, the genetic correlations between colour and the liveweights are low to medium.

The genetic and phenotypic correlations between pairs of the fleece traits are more variable than the correlations between the liveweight traits, and there is less consistency in terms of sign and size of the genetic correlation relative to the phenotypic correlation.

Lamb fleece weight has a high genetic and a medium phenotypic correlation estimate with hogget fleece weight. It has already been noted that hogget fleece weight is a very important selection criteria in the Romney and a reasonable positive phenotypic correlation means early selection of lambs using fleece weight records may be of benefit. However, given the stronger correlation between weaning and later liveweights, and the importance of this trait also as a selection criterion

in the Romney; preliminary selection based on liveweight is probably more efficient. The positive correlations between liveweight and fleece weight are important in this context.

There are few previous estimates available of correlations between lamb fleece weight and other fleece traits in the Romney. The genetic correlation of 0.80 between lamb and hogget fleece weight obtained by Baker (1977) is larger than the estimate of 0.59 obtained here, although both represent a high correlation. The estimates of the correlations between lamb fleece weight and fleece quality traits measured on the hogget fleece in Table 5.2 are variable in size, but generally low and in the expected direction.

The genetic correlation estimates in Table 5.2 of hogget fleece weight with the other fleece traits are high for staple length, character and break severity; medium for quality number and low for colour. The estimates in Table 2.5 are generally medium to high for staple length and quality number, but are quite variable for the other subjective traits. The estimates obtained in this study of the genetic correlation between hogget fleece weight and other fleece traits are therefore in agreement with the common range of estimates obtained in earlier studies.

The phenotypic correlations between hogget fleece weight and the other hogget fleece traits are generally lower than the genetic correlations, but agree in terms of sign. An exception is the phenotypic correlation between hogget fleece weight and colour which is slightly larger and positive compared to the low, negative genetic correlation. There is general agreement also between the estimates of this study and the corresponding estimates in Table 2.4.

The correlations between fleece weight and quality number could be antagonistic if a reduction in fibre diameter is a selection objective. However, this is not always an objective in Romney flocks; it is usually sufficient to just maintain the average fibre diameter of the flock below a certain "ceiling" level.

Character is not now regarded as an important selection criterion, therefore, the low variable correlations associated with it are of less importance. Since colour can be important in the Romney, as noted earlier, the negative genetic correlation between hogget fleece weight and colour found in this study, which is supported by the few estimates in Table 2.5, is undesirable, although it is small.

The estimates of the genetic and phenotypic correlations between staple length and the fleece quality traits are consistent in terms of sign and similar in terms of size. The estimates fall within the common range of estimates reviewed in Tables 2.4 and 2.5. Staple length, in itself, is not a very important selection criterion, although short fibres in the staple may cause processing problems; for example, in worsted yarn manufacture (Wickham, 1982a). Increased staple length is, as noted before, positively correlated with an increase in fleece weight which is an advantage. The negative correlation between staple length and quality number found in this study is lower than the estimates, both genetic and phenotypic, previously obtained. Negative correlations between fleece quantity traits and quality number are antagonistic in most Romney breeding-ewe flocks and fibre diameter needs to be monitored accordingly. However, depending on the end-product (e.g., carpet wools) they may, in fact, be used to advantage. Generally speaking, however, it is important in the Romney to keep a check on fleece quality traits when selecting for increased fleece weight.

CHAPTER SEVEN

CONCLUSION

## CHAPTER SEVEN

CONCLUSION

This study was concerned with the estimation of variance components using REML; a new method of estimation with respect to analysing sheep breeding data in New Zealand. The estimates of the variance components were then used to generate estimates of the heritabilities, phenotypic and genetic correlations, and the BLUE's of the non-genetic (fixed) effects.

In a dual-purpose sheep breeds, such as the Romney, where both meat and wool production are important selection objectives, a selection index approach is an efficient method of genetic improvement. Accordingly, Sheepplan uses a selection index for dual-purpose breeds which incorporates number of lambs born, weaning weight, spring (hogget) liveweight and hogget fleece weight (Clarke and Rae, 1976).

In most cases, the parameter estimates from this study were in general agreement with the results from previous studies and also generally agreed with the parameter values currently used by Sheepplan. However, the  $h^2$  estimates of the liveweight traits were found to be lower than those used by Sheepplan, although they were relatively comparable with the more recent  $h^2$  estimates of liveweight in other studies. Genetic parameters may change with time; hence, it is possible that the  $h^2$  of liveweight has decreased in the Romney. If  $h^2$  of liveweight is, in fact, lower than the Sheepplan parameters, then current selection response for liveweight may not be as rapid as would be expected from Sheepplan.

The genetic correlation estimates between WWT and the liveweights

from JUN to 2TH were also found to be lower than the values used by Sheeplan. However, these results have not been supported by other studies. The genetic correlation between WWT and NOV (hogget) live-weight is of particular interest since it is involved in the Sheeplan selection index.

With the exception of these differences, it is interesting to note that the REML parameter estimates of this study were generally similar to the results of previous studies where other estimation methods (notably Henderson's Methods; Henderson, 1953) have been used. The close similarity, in this study, of the REML  $h^2$  estimates with the Method III  $h^2$  estimates is also of interest, especially since Method III has the desirable feature of unbiasedness, which REML does not.

The efficiency with which a selection index actually identifies the individuals of higher genetic merit, is affected by the reliability of the parameters included. Thus, accurate data collection and reliable parameter estimates are important. Genetic and environmental changes in the population, technological improvements which improve data collection and processing and the development of new estimation methods with specific desirable properties, makes it worthwhile to periodically re-estimate the genetic and phenotypic parameters of the population.

BIBLIOGRAPHY

BIBLIOGRAPHY

- Adby, P.R. and Dempster, M.A.H. (1974). Introduction to Optimization Methods. Chapman and Hall, London.
- Albert, A. (1976). When is a sum of squares an analysis of variance? Ann. Stat., 4: 775-778.
- Anderson, R.D. (1978). Studies on the estimation of variance components. Ph.D. Thesis, Cornell University, Ithaca, New York.
- Anderson, R.D. (1979). Estimating variance components from balanced data: optimum properties of REML solutions and MIVQUE estimators. In Variance Components and Animal Breeding. Proc. of a Conference in Honor of C.R. Henderson held at Cornell University, Ithaca, New York.
- Anderson, R.D., Henderson, H.V., Pukelsheim, F. and Searle, S.R. (1979). Best estimation of variance components from balanced data, with arbitrary kurtosis. (Submitted).
- Anderson, R.L. and Bancroft, T.A. (1952). Statistical Theory in Research. McGraw-Hill Book Co., Inc., New York.
- Anderson, T.W. (1973). Asymptotically efficient estimation of covariance matrices with linear structure. Ann. Stat., 1: 135-141.
- Anderson, T.W. (1969). Statistical inference for covariance matrices with linear structure. In Multivariate Analysis - II. P.R. Krishnaiah (Ed.). Academic Press, New York.
- Baker, R.L. (1974). Breeding for improved productivity in Romney sheep - Waikite. Anim. Breed. Abstr., 45: Abstr. No. 541.
- Baker, R.L. (1977). Genetic and phenotypic parameters in Romney sheep. Anim. Breed. Abstr., 48: Abstr. No. 3744.

- Baker, R.L., Clarke, J.N. and Carter, A.H. (1974). Sources of variation for wool body weight and oestrus characters in Romney hoggets. Proc. N.Z. Soc. Anim. Prod., 34: 19-22.
- Baker, R.L., Clarke, J.N., Carter, A.H. and Diprose, G.D., (1979). Genetic and phenotypic parameters in New Zealand Romney sheep. I. Body weights, fleece weights and oestrous activity. N.Z. J. agric. Res., 22: 9-21.
- Bard, Y. (1974). Nonlinear Parameter Estimation. Academic Press, New York.
- Blair, H.T. (1981). Response to selection for open face and greasy fleece weight in Romney sheep. Ph.D. Thesis, Massey University.
- Brown, G.H. and Turner, H.N. (1968). Response to selection in Australian Merino sheep. II. Estimates of phenotypic and genetic parameters for some production traits in Merino ewes and an analysis of the possible effects of selection on them. Aust. J. agric. Res., 19: 303-322.
- Brown, K.G. (1976). Asymptotic behaviour of MINQUE-type estimators of variance components. Ann. Stat., 4: 746-754.
- Callow, C. and Johnson, D.L. (1982). Technical developments to the Sheeplan (New Zealand's National Flock Recording Scheme) selection indices. (Unpubl.).
- Carter, R.L. (1981). Restricted maximum likelihood estimation of bias and reliability in the comparison of several measuring methods. Biometrics, 37: 733-741.
- Ch'ang, T.S. and Rae, A.L. (1961). Sources of variation in the weaning weight of Romney Marsh lambs. N.Z. J. agric. Res., 4: 478-482.

- Ch'ang, T.S. and Rae, A.L. (1970). The genetic basis of growth, reproduction and maternal environment in Romney ewes. I. Genetic variation in hogget characters and fertility of the ewe. Aust. J. agric. Res., 21: 115-129.
- Ch'ang, T.S. and Rae, A.L. (1972). The genetic basis of growth, reproduction and maternal environment in Romney ewes. II. Genetic covariation between hogget characters, fertility and maternal environment of the ewe. Aust. J. agric. Res., 23: 149-165.
- Chopra, S.C. (1978). Genotype-environment interactions and genetic parameters in New Zealand Romney sheep. Ph.D. Thesis, Massey University.
- Clarke, J.N. and Rae, A.L. (1976). Sheeplan Advisors Manual. Ministry of Agriculture and Fisheries, Wellington.
- Cochran, W.G. (1939). The use of the analysis of variance in enumeration by sampling. J. Amer. Stat. Ass., 34: 492-510.
- Corbeil, R.R. and Searle, S.R. (1976a). Restricted Maximum Likelihood (REML) estimation of variance components in the mixed model. Technometrics, 18: 31-38.
- Corbeil, R.R. and Searle, S.R. (1976b). A comparison of variance components estimators. Biometrics, 32: 779-791.
- Crump, S.L. (1947). The estimation of components of variance in multiple classifications. Ph.D. Thesis, Iowa State University, Ames, Iowa.
- Dalton, D.C. and Rae, A.L. (1978). The New Zealand Romney sheep: A review of productive performance. Anim. Breed. Abstr., 46: 657-680.

- Eikje, E.D. (1978a). Comparison of adjustment methods for lamb weaning weight. (Unpubl. paper).
- Eikje, E.D. (1978b). Genetic parameters for lamb growth rate, hogget fleece weight and quality number in the New Zealand Romney. (Unpubl. paper).
- Eikje, E.D. (1979). Report on research with NRAC - Post-Doctoral Fellowship from July 1976 to June 1978. (Unpubl. paper).
- Eikje, E.D. and Johnson D.L. (1979). Adjustment factors for lamb weaning weight. II. Sources of variation in adjustment factors. N.Z. J. agric. Res., 22: 391-397.
- Eisenhart, C. (1947). The assumptions underlying the analysis of variance. Biometrics, 3: 1-21.
- Fisher, R.A. (1918). The correlation between relatives of the supposition of mendelian inheritance. Trans. Roy. Soc. Edin., 52: 399-433.
- Fisher, R.A. (1925). Statistical Methods for Research Workers. 12th revised edition (Ed. F.A.E. Crew). Oliver and Boyd, Edinburgh.
- Fisher, R.A. (1930). The Genetical Theory of Natural Selection. Clarendon Press, Oxford, United Kingdom.
- Ganguli, M. (1941). A note on nested sampling. Sankhyā, 5: 449-452.
- Gill, P.E. and Murray, W.A. (1974)(Eds). Numerical Methods for Constrained Optimization. Academic Press, New York.
- Goodnight, J.H. (1976). Maximally-invariant quadratic unbiased estimators. Biometrics, 32: 477-480.
- Graybill, F.A. (1954). On quadratic estimates of variance components. Ann. Math. Stat., 25: 367-372.

- Graybill, F.A. and Hultquist, R.A. (1961). Some theorems concerning Eisenhart's Model II. Ann. Math. Stat., 32: 261-269.
- Graybill, F.A. and Wortham, A.W. (1956). A note on uniformly best unbiased estimators for variance components. J. Amer. Stat. Ass., 51: 266-268.
- Haldane, J.B.S. (1946). The interaction of nature and nurture. Ann. Eugenics (Lond.), 13: 197.
- Hartley, H.O. and Rao, J.N.K. (1967). Maximum likelihood estimation for the mixed analysis of variance model. Biometrika, 54: 93-108.
- Hartley, H.O., Rao, J.N.K. and LaMotte, L.R. (1978). A simple "synthesis"-based method of variance component estimation. Biometrics, 34: 233-242.
- Harville, D.A. (1969). Quadratic unbiased estimation of variance components for the one-way classification. Biometrika, 56: 313-326.
- Harville, D.A. (1975). Index selection with proportionality constraints. Biometrics, 31: 223-225.
- Harville, D.A. (1976). Extension of the Gauss-Markov theorem to include the estimation of random effects. Ann. Stat., 4: 384-395.
- Harville, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. J. Amer. Stat. Ass., 72: 320-340.
- Harville, D.A. (1981). Unbiased and minimum variance unbiased estimation of estimable functions for fixed linear models with arbitrary covariance structure. Ann. Stat., 9(3): 633-637.

- Hazel, L.N. (1943). The genetic basis for constructing selection indexes. Genetics, 28: 476-490.
- Hemmerle, W.J. and Hartley, H.O. (1973). Computing maximum likelihood estimates for the mixed A.O.V. model using the W-transformation. Technometrics, 15: 819-832.
- Hemmerle, W.J. and Lorens, J.A. (1976). Improved algorithm for the W-transform in variance component estimation. Technometrics, 18: 207-212.
- Henderson, C.R. (1953). Estimation of variance and covariance components. Biometrics, 9: 226-252.
- Henderson, C.R. (1963). Selection index and expected genetic advance. Statistical Genetics and Plant Breeding Symposium National Research Publication 982, held at North Carolina State College, Raleigh, North Carolina, 20-29 March 1961. (Eds W.D. Hanson and H.F. Robinson).
- Henderson, C.R. (1972a). Sire evaluation and genetic trends. Proc. of the Animal Breeding and Genetics Symposium in honor of Dr J.L. Lush : 10-41.
- Henderson, C.R. (1972b). Prediction. (Unpubl. notes).
- Henderson, C.R. (1973). Maximum likelihood estimation of variance components. (Unpubl. paper).
- Henderson, C.R. (1980). A simple method for unbiased estimation of variance components in the mixed model. Presented at the 72nd Annual Meeting of the American Society of Animal Science, held at Cornell University, Ithaca, New York.
- Henderson, C.R., Kempthorne, O., Searle, S.R. and Von Krosigk, C.M. (1959). The estimation of environmental and genetic trends from records subject to culling. Biometrics, 15: 192-218.

- Henderson, C.R. and Quaas, R.L. (1977). A comparison of method 3 and maximum likelihood estimates of heritability and genetic correlations. Mimeo. paper presented at the Annual Meeting of the American Society of Animal Science.
- Henderson, C.R., Searle, S.R. and Schaeffer, L.R. (1974). The invariance and calculation of method 2 for estimating variance components. Biometrics, 30: 583-588.
- Henderson, C.R., Ufford, G.R. and Schaeffer, L.R. (1976). Maximum likelihood estimation of the parameters and the conditional means of the random variables in the mixed model. Mimeo. paper in the Biometrics Unit, Cornell University, Ithaca, New York.
- Hight, G.K. and Jury, K.E. (1970). Hill country sheep production. II. Lamb mortality and birth weights in Romney and Border Leicester x Romney flocks. N.Z. J. agric. Res., 13: 735-752.
- Hocking, R.R. and Kutner, M.H. (1975). Some analytical and numerical comparisons of estimators for the mixed A.O.V. model. Biometrics, 31: 19-28.
- Hudson, G.F.S. and Van Vleck, L.D. (1982). Estimation of components of variance by method 3 and Henderson's new method. J. Dairy Sci., 65: 435-441.
- Hultquist, R.A. and Graybill, F.A. (1965). Minimum sufficient statistics for the two-way classification mixed model design. J. Amer. Stat. Ass., 60: 182-192.
- Jennrich, R.I. and Sampson, P.F. (1976). Newton-Raphson and related algorithms for maximum likelihood estimation of variance components. Technometrics, 18: 11-18.

- Jury, K.E., Johnson, D.L. and Clarke, J.N. (1979). Adjustment factors for lamb weaning weight. I. Estimates from commercial flocks. N.Z. J. agric. Res., 22: 385-389.
- LaMotte, L.R. (1970). A class of estimators of variance components. Technical Report No. 10. Department of Statistics, University of Kentucky, Lexington, Kentucky.
- LaMotte, L.R. (1973a). Quadratic estimation of variance components. Biometrics, 29: 311-330.
- LaMotte, L.R. (1973b). On non-negative quadratic unbiased estimation of variance components. J. Amer. Stat. Ass., 68: 728-730.
- Lundie, R.S. (1971). Studies of productive traits in a New Zealand Romney flock: The effect of some environmental factors, heritabilities and repeatabilities. M.Agr.Sci. Thesis, Massey University.
- Lush, J.L. (1937). Animal Breeding Plans. Iowa State College Press, Iowa.
- McCutcheon, S.N. (1981). A study of some factors affecting the resistance of newborn lambs to cold-stress with particular reference to starvation and exposure mortality. Ph.D. Thesis, Massey University.
- McMahon, P.R. (1943). The inheritance of multifactor characters in sheep. Proc. N.Z. Soc. Anim. Prod., 3: 70-80.
- McPherson, A.W. (1982). An appraisal of selection objectives and criteria for New Zealand Romney sheep with particular reference to wool traits. M.Agr.Sci. Thesis, Massey University.

- Miller, J.J. (1973). Asymtotic properties of maximum likelihood estimates in the mixed model of the analysis of variance. Technical Report No. 12, Department of Statistics, Stanford University, Stanford, California.
- Miller, J.J. (1977). Asymtotic properties of maximum likelihood estimates in the mixed model of the analysis of variance. Ann. Stat., 5: 746-762.
- Ministry of Agriculture and Fisheries, Agricultural Research Division. Annual Report, 1970/71, Wellington.
- Ministry of Agriculture and Fisheries, Agricultural Research Division. Annual Report, 1971/72, Wellington.
- Ministry of Agriculture and Fisheries, Agricultural Research Division. Annual Report, 1972/73, Wellington.
- Mood, A.M., Graybill, F.A. and Boes, D.C. (1974). Introduction to the Theory of Statistics. 3rd ed. McGraw-Hill, Inc., Tokyo.
- Mullaney, P.D., Brown, G.H., Young, S.S.Y. and Hyland, P.G. (1970). Genetic and phenotypic parameters for wool characteristics in fine-wool Merino, Corriedale and Polwarth sheep. II. Phenotypic and genetic correlations, heritability and repeatability. Aust. J. agric. Res., 21: 527-540.
- Nawara, W. and Duniec, H. (1972). Heritability of the main production traits in Polish Mountain sheep. Anim. Breed. Abstr., 41: Abstr. No. 2144.
- Patterson, H.D. and Thompson, R. (1971). Recovery of interblock information when block sizes are unequal. Biometrika, 58: 545-554.

- Patterson, H.D. and Thompson, R. (1974). Maximum likelihood estimation of components of variance. Proc. of the 8th International Biometric Conference: 197-207.
- Powell, M.J.D. (1970). A survey of numerical methods for unconstrained optimization. S.I.A.M. Review, 12: 79-97.
- Quaas, R.L. (1976). What is restricted maximum likelihood? Paper presented at the October 1976 Animal Breeding Seminar, Cornell University, Ithaca, New York.
- Quaas, R.L. and Bolgiano, D.C. (1979). Sampling variances of the MIVQUE and method 3 estimators of the sire component of variance. In Variance Components and Animal Breeding. Proc. of a Conference in honor of C.R. Henderson, held at Cornell University, Ithaca, New York.
- Rae, A.L. (1946). Some aspects of progeny testing of New Zealand Romney Marsh rams. M.Agr.Sci. Thesis, Massey University.
- Rae, A.L. (1948). Some aspects of progeny testing in sheep. Proc. N.Z. Soc. Anim. Prod., 8: 97-104.
- Rae, A.L. (1950). Genetic variation and covariation in productive traits of New Zealand Romney Marsh sheep. Ph.D. Thesis, Iowa State College, Iowa.
- Rae, A.L. (1958a). Genetic variation and covariation in productive characteristics of New Zealand Romney Marsh sheep. N.Z. J. agric. Res., 1: 104-123.
- Rae, A.L. (1977b). (Pers. comm.).
- Radomska, M.J. and Tyszka, A. (1972). I. Phenotypic and genetic correlations of some traits in Romney Marsh sheep. Anim. Breed. Abstr., 41: Abstr. No. 1137.

- Radomska, M.J. and Klewicz, J. (1975). A comparative analysis of certain Romney Marsh flocks. II. Genetic parameters. Prace I. Materialy Zootechniczne, 7: 35-40.
- Radomska, M.J. and Klewicz, J. (1976). A comparative analysis of certain Romney Marsh flocks. III. Selection indices. Prace I. Materialy Zootechniczne, 11: 17-23.
- Rao, C.R. (1970). Estimation of heteroscedastic variances in linear models. J. Amer. Stat. Ass., 65: 161-172.
- Rao, C.R. (1971a). Estimation of variance and covariance components - MINQUE theory. J. Multivar. Anal., 1: 257-275.
- Rao, C.R. (1971b). Minimum variance quadratic unbiased estimation of variance components. J. Multivar. Anal., 1: 445-456.
- Rao, C.R. (1972). Estimation of variance and covariance components in linear models. J. Amer. Stat. Ass., 67: 112-115.
- Rao, C.R. (1973). Linear Statistical Inference and its Applications. 2nd ed. John Wiley and Sons, New York.
- Rao, C.R. (1977). Discussion of Harville (1977). J. Amer. Stat. Ass., 72: 320-340.
- Rao, C.R. (1979). Estimation of parameters in the singular Gauss-Markoff model. Commun. Statist., A8: 1353-1358.
- Rao, P.S.R.S. (1977). Theory of the MINQUE - A review. Sankhyā, 39 (Series B): 201-210.
- Ross, D.A. (1964). The relation of count to other characteristics of New Zealand wools. N.Z. J. agric. Res., 7: 666-667.
- Rothschild, M.F. and Henderson, C.R. (1979). Maximum likelihood estimates of parameters of first and second lactation milk records. J. Dairy Sci., 62: 990-995.

- Rothschild, M.F. Henderson, C.R. and Quaas, R.L. (1979). Effects of selection on variances and covariances of simulated first and second lactations. J. Dairy Sci., 62: 996.
- Searle, S.R. (1968). Another look at Henderson's methods of estimating variance components. Biometrics, 24: 749-787.
- Searle, S.R. (1970). Large sample variances of maximum likelihood estimators of variance components using unbalanced data. Biometrics, 27: 1-76.
- Searle, S.R. (1971a). Linear Models. John Wiley and Sons, New York.
- Searle, S.R. (1971b). Topics in variance component estimation. Biometrics, 27: 1-76.
- Searle, S.R. (1977). Variance components estimation: A thumbnail review. (Unpubl. paper).
- Searle, S.R. (1979a). Notes on variance component estimation. A detailed account of maximum likelihood and kindred methodology. (Unpubl. paper). Biometrics Unit, Cornell University, Ithaca, New York.
- Searle, S.R. (1979b). Maximum likelihood and minimum variance estimation of variance components. In Variance Components and Animal Breeding. Proc. of a Conference in honor of C.R. Henderson, held at Cornell University, Ithaca, New York.
- Schaeffer, L.R. (1979). Estimation of variance and covariance components for average daily gain and backfat thickness in swine. In Variance Components and Animal Breeding. Proc. of a Conference in honor of C.R. Henderson, held at Cornell University, Ithaca, New York.

- Schaeffer, L.R. (1981). Estimation of genetic covariances from sire progeny test results. J. Dairy Sci., 64: 546-551.
- Seely, J. (1971). Quadratic subspaces and completeness. Ann. Math. Stat., 42: 710-721.
- Sumner, R.M.W. (1969). A comparative study of the effect of two stocking levels on wool follicle development and wool production of the New Zealand Romney. M.Agr.Sci. Thesis, Massey University.
- Szatrowski, T.H. and Miller, J.J. (1977). Explicit maximum likelihood estimates in the mixed model of the analysis of variance. (Unpubl. mimeo. paper). Department of Statistics, Rutgers University, New Brunswick, New Jersey.
- Thompson, W.A. (1962). The problem of negative estimates of variance components. Ann. Math. Stat., 33: 273-289.
- Thompson, R. (1969). Iterative estimation of variance components for non-orthogonal data. Biometrics, 25: 767-773.
- Tippet, L.H.C. (1931). The Methods of Statistics. 1st ed. Williams and Norgate Ltd, London.
- Townsend, E.C. (1968). Unbiased estimators of variance components in simple unbalanced designs. Ph.D. Thesis, Cornell University, Ithaca, New York.
- Townsend, E.C. and Searle, S.R. (1971). Best quadratic unbiased estimation of variance components from unbalanced data in the 1-way classification. Biometrics, 27: 643-657.
- Tripathy, L.F. (1966). Estimation of important genetic and phenotypic parameters of some productive traits in the New Zealand Romney Marsh sheep. M.Agr.Sci. Thesis, Massey University.

- Turner, H.N. and Young, S.S.Y. (1969). Quantitative Genetics in Sheep Breeding. McMillan, Australia.
- Ugalde, O.J. (1978). Analysis of some genetic and environmental factors affecting birth weight and growth to three months in Romney Marsh sheep. Anim. Breed. Abstr., 47: Abstr. No. 5493.
- Wickham, G.A., et al. (Unpubl.). Cited by Sumner (1969).
- Wickham, G.A. (1982a). New Zealand crossbred wool: Aspects of its production uses and prospects. Proc. N.Z. Soc. Anim. Prod., 42: 1-5.
- Wickham, G.A. (1982b). Selection criteria for young animals being mated early in life and genetic implications of early breeding. N.Z. Agric. Sci., 16: 110-112.
- Winsor, C.P. and Clarke, G.L. (1940). A statistical study of variation in the catch of plankton nets. Sears Foundation J. Marine Res., 3: 1-27.
- Wright, G.M. and Stevens, P.G. (1953). Lifetime wool production and breeding performance of Romney Marsh and Corriedale ewes. N.Z. J. Sci. and Tech. (A), 34: 430-435.
- Wright, S. (1939). Genetic principles governing the rate of progress of livestock breeding. Proc. Amer. Soc. Anim. Prod., 1938-39: 18-26.
- Yates, F. (1934). The analysis of multiple classifications with unequal numbers in the different subclasses. J. Amer. Stat. Ass., 29: 51-66.

- Young, S.S.Y., Turner, H.N. and Dolling, C.H.S. (1960). Comparison of estimates of repeatability and heritability for some production traits in Merino rams and ewes. II. Heritability. Aust. J. agric. Res., 11: 604-617.
- Zelen, M. (1968). Discussion of Searle (1968). Biometrics, 24: 779-780.

APPENDIX

## APPENDIX

Obtaining  $\tilde{T}$  from the Mixed Model Equations (MME)

Solving Henderson's MME, as part of the REML algorithm discussed in Chapter Four, yields the inverse of the left-hand-side matrix, namely:

$$\tilde{A}^{-1} = \begin{bmatrix} \tilde{X}'\tilde{X} & \tilde{X}'\tilde{Z} \\ \tilde{Z}'\tilde{X} & \tilde{Z}'\tilde{Z} + \tilde{I}\sigma_e^2/\sigma_s^2 \end{bmatrix}^{-1} = \begin{bmatrix} \tilde{K} & \tilde{L} \\ \tilde{L}' & \tilde{M} \end{bmatrix}$$

where  $\tilde{A}$  is a symmetric matrix with full rank under the subclass means model of (4.2).

Following Searle (1968), the right-hand submatrix of  $\tilde{A}^{-1}$  will be:

$$\tilde{M} = [\tilde{Z}'\tilde{Z} + \tilde{I}\sigma_e^2/\sigma_s^2 - \tilde{Z}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Z}]^{-1}$$

Now,

for  $\tilde{D} = \tilde{I}\hat{\sigma}_s^2$  from (4.2) and  $\tilde{T} = (\tilde{I} + \tilde{Z}'\tilde{S}\tilde{Z}\tilde{D})^{-1}$  of (2.21),

$$\tilde{D}\tilde{T} = (\tilde{D}^{-1} + \tilde{Z}'\tilde{S}\tilde{Z})^{-1}$$

and,

for  $\tilde{S} = \tilde{R}^{-1} - \tilde{R}^{-1}\tilde{X}(\tilde{X}'\tilde{R}^{-1}\tilde{X})^{-1}\tilde{X}\tilde{R}^{-1}$

where  $\tilde{R} = \tilde{I}\hat{\sigma}_e^2$  from (4.2)

it is readily shown that,

$$\tilde{Z}'\tilde{S}\tilde{Z} = 1/\hat{\sigma}_e^2 [\tilde{Z}'\tilde{Z} - \tilde{Z}'\tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Z}]$$

so that,

$$\begin{aligned} \tilde{M} &= [\tilde{D}^{-1}\hat{\sigma}_e^2 + \tilde{Z}'\tilde{S}\tilde{Z}\hat{\sigma}_e^2]^{-1} \\ &= \tilde{D}\hat{\tilde{T}}1/\hat{\sigma}_e^2 \end{aligned}$$

That is,

the lower right-hand submatrix of  $\hat{\tilde{A}}^{-1}$  contains  $\hat{\tilde{T}}\hat{\sigma}_s^2/\hat{\sigma}_e^2$  and therefore multiplying the elements of this submatrix by  $\hat{\sigma}_e^2/\hat{\sigma}_s^2$  yields  $\hat{\tilde{T}}$ .