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.

# A STUDY OF BIASES IN DAIRY
# SIRE EVALUATION

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

ROBERT DONALD ANDERSON

1974

# ABSTRACT

A total of 8,684 first-lactation records produced in the Auckland region in the dairying seasons 1966-67 to 1972-73, inclusive were assembled. Based on weighted and unweighted intra-sire regressions of progeny performance (deviation from within-[illegible] and within-year contemporary average) on time, a series of estimates of the genetic trend in production for the period were obtained. Variable estimates, allied with large standard errors [illegible] precluded reliable conclusions on the nature of the genetic [illegible]

By applying a mixed model solution method of sire-evaluation [illegible] but only to a restricted sub-sample of the data (2,155 records) [illegible] breeding values of 47 sires were estimated unbiased by [illegible] The numbers of records involved were too few to allow a [illegible] comparison of these estimates with estimates of the [illegible] of the same sires obtained by the Farm Production D[illegible] N.Z. Dairy Board using its method of sire evaluation

Based on variance component estimates obtained [illegible] of Henderson's Method I, heritability of milk yield and [illegible] was estimated to be 0.43 and 0.36 respectively. These [illegible] agree well with estimates reported elsewhere.

Additional estimates of the genetic trend were obtained [illegible] regressing the solutions for the fixed effects (sire-group effects of the sire evaluation model on time.

# ACKNOWLEDGEMENTS

The author is specially indebted to his supervisor, Professor A. L. Rae, for wise and invaluable guidance and assistance in all aspects of the study.

Sincere thanks are due to his colleague, Mr R.A. Barton, for helpful criticism of the manuscript and for advice on many occasions.

Grateful acknowledgement is made to Dr C. R. Boswell Dr G.A. Wickham and Messrs C. L. Read and D. J. Wilson for their assistance and guidance in computing matters.

Gratitude is extended to Mr J. W. Stichbury (N. Z. Da for his making available the data used in this study and to Miss V. Brickwedde for help in assembling the data.

It is a pleasure to acknowledge Mrs H. Horsfall for her skilled typing of the manuscript.

Finally, very special thanks are due to his wife, Ro her understanding, support and encouragement throughout

# TABLE OF CONTENTS

# LIST OF TABLES

# CHAPTER 1

# INTRODUCTION

The commercial development of artificial insemination [illegible] ) in the New Zealand dairy industry resulted from the realiz[illegible] at, through the use of sires of superior genetic merit, it prov[illegible] means for increasing the rate of genetic gain in cattle popu[illegible]. In recent years, the industry has witnessed a dramatic increase [illegible] the numbers of inseminations per proven sire. However if the genetic gains resulting from the widespread use of sires through A.I. are to be maximized, it is imperative that the sires [illegible] accurately evaluated. It is ironical, that the major [illegible] an A.I. programme, that is, elevation of the genetic [illegible] population, has been demonstrated elsewhere to be an imp[illegible] source of bias in the evaluation of dairy sires.

The methods of sire evaluation that are susceptible to the bias due to genetic trend are those which rely on comparisons of the average production of the daughters of individual sires with the average production of representatives of the population [illegible] herdmates in a series of herds.

If the genetic merit of the overall population is increas[illegible] since the genotype of a sire is constant, the productio[illegible] successive groups of daughters of the sire (of the sa[illegible] their herdmates, will decline with time. A clear demon[illegible] this effect was given by the study of Carter (1969). In [illegible] investigation, 19 young bulls were used through A.I. and they [illegible] aside for a period of five years before being returned to A.I. service. The average production superiority of their daughters sired at the end of the five-year period was 260 lb. of milk less than the production superiority of their half sisters sired five years earlier. The question raised by this effect is: "For comparative purposes, at which point was the valid estimate of the genetic merit of these bulls obtained?" Obviously, the appropriate estimate is that which is most recent, but the point requiring consideration, is whether any account is taken of the earlier appraisals of these sires.

If the estimates of the genetic merit of all sires are based on the productive performance of a restricted group of daughters (for instance, the first-lactation records made in the most recent dairying season by 2-year-old heifers), genetic trend cannot bias sire comparisons. However, in order that precision of estimation be maximized, sire evaluation procedures normally utilize the lactation records of as many daughters as possible and therefore the estimate of a sire's genetic merit may be based on lactation records produced in a number of dairying seasons. Under these conditions, if a sire whose genetic merit is based on the performance of daughters in a number of seasons is compared to a sire being sampled for the first time, bias due to genetic trend will occur since the older sire has the advantage of having had some of its daughters compared to the population average when it was at a lower level. In other words, genetic trend in the population has the effect of favouring long-standing sires when their estimated genetic merit is being compared to that of a sire recently entering service.

The objective of the present study was to determine whether genetic trend is an important source of bias in the method of dairy sire evaluation currently operated by the N.Z. Dairy Board. In this respect, the study involved two broad aspects. These were:

(i) Estimating the genetic trend currently occurring in the dairy cattle population in New Zealand.

(ii) Estimating the genetic merit of a series of sires using a method of evaluation known to be capable of ranking sires unbiased by genetic trend and, to assess whether the sire rankings obtained agreed with the rankings produced by the N.Z. Dairy Board's method of evaluation.

# CHAPTER 2

# REVIEW OF LITERATURE

## I. THE IMPORTANCE OF ACCURATE SIRE EVALUATION

The need for accurate evaluation of the genetic merit of sires used for A.I. in New Zealand has increased in recent years. Advances in the technology of semen dilution (Shannon, 1968) have allowed a bull to be mated to many thousands of cows. For example, in the A.I. scheme operated by the N.Z. Dairy Board, the maximum number of inseminations for a bull in a single mating season has risen from 10,404 in 1963 (N.Z. Dairy Board, 1972) to [illegible] in 1973 (Macmillan, <u>pers. comm.</u>). Clifford (1972) has estimated that a bull with a rating[a] of +50 lb. milkfat, having sired 50,000 daughters, would be responsible for 12,500,000 lb. more milkfat than [illegible] rating was zero. Based upon the gauranteed milkfat p[illegible] 50 cents per lb. for the 1973 season, the extra production would amount to $6,250,000. The genetic implications for New Zealand of widespread usage of proven sires has been discussed by Stichbury (1968).

Large-scale usage of proven bulls for A.I. reflects the importance of a sire's genetic contribution to a subsequent generation, a feature which has long been recognized. Ward (1937) stated: "Undoubtedly as is now generally admitted, the herd sire is the greatest [illegible] influence in raising or lowering the production level of the herd

---

[a] A bull's rating is the expected average production deviation of his future daughters from the population average.

Robertson and Rendel (1950) quantitatively apportioned the genetic improvement that may result, under optimal conditions, for a breeding unit of 2,000 cows, an exercise which has been repeated by Legates (1971) for a breeding unit of 60,000 cows. Table 2.1 summarizes the results of these studies and demonstrates that the genetic contribution of the sire, relative to that of the dam, is more important.

The results obtained by Legates are of special interest in that the breeding unit concerned, in terms of cow numbers (60,000), is more typical of an artificially-bred population than is that of Robertson and Rendel (2,000).

Studies conducted by the N.Z. Dairy Board (1943, 1961), Macmillan (1973), Macmillan and Watson (1973) demonstrated an important antagonism between herd wastage rate and the number of replacement heifers available for selection under average New Zealand dairyfarming conditions. The annual wastage rate in New Zealand is 21 percent (N.Z. Dairy Board, 1972) which makes it difficult for a farmer to exert any selection pressure in the choice of dams to produce daughters. Consequently, the contribution of this pathway to improvement is small. Studies by Jackson (1973) demonstrated that the overall contribution of the dam to improvement may be strengthened through the dam to son pathway, if a system of contract mating is linked with an A.I. programme.

The ideas presented in this section support the viewpoint that production gains, resulting from selection, will be maximized by extensive use of sires which have been accurately assessed to be of superior genetic merit.

Table 2.1 : Relative contribution of the four genetic pathways to genetic gain.

| Genetic pathway | Estimated contribution to improvement (%) | |
|---|---|---|
| | Robertson and Rendel (1950) | Legates (1971) |
| Sire to son | 43 | 46 |
| Sire to daughter | 18 | 23 |
| Dam to son | 33 | 23 |
| Dam to daughter | 6 | 8 |

## II. SIRE EVALUATION IN NEW ZEALAND UP TO 1969

The evolution of the N.Z. Dairy Board's method of sire evaluation up to 1962 has been reviewed by Edey (1966). Whilst providing valuable historical information, Edey's review lacks mathematical detail, particularly insofar as the method in use by 1962 is concerned. At that time the genetic merit of a sire was estimated using the "difference from expectancy" method of evaluation. A detailed description of this method was provided by Searle (1964a) when comparing the New Zealand method of sire evaluation with those of Great Britain and New York State. According to Searle, the genetic merit of an A.I. sire was estimated using the following procedure:

$$\text{Estimated sire merit} = 2\ \text{rating} + \text{breed average}$$

$$\text{where: Rating} = F\left[\ \overline{X} - (B + 0.9\,(\overline{S} - B)\,)\right]$$
$$= F\left[\ \overline{X} - 0.9\overline{S} - 0.1B\right]$$

and where: $B$ = The breed average of all cows in the particular Herd Improvement Association.

$\overline{S}$ = The overall mean of the herdmate averages to which the daughters of the sire were compared.

$\overline{X}$ = The average production of the daughters of the sire being evaluated.

$0.9$ = The regression of daughter average on true herd average (effectively intra-sire regression of daughter average on herdmate average).

$F$ = The regression of the sire's future daughter production on his estimated true daughter average.

There are a number of features of this method of evaluation which warrant discussion:

## A. Age-Correction Factors

All averages (breed, herdmate and daughter) were calculated on a "mature equivalent" basis to remove the effect of age at calving on production. Animals aged 5 to 9 years inclusive, were considered to be mature. Those aged 2, 3 and 4 years had their production records age corrected to a mature cow level using simplified herd-level age-correction factors obtained from the method of Searle (1960). Cows aged 10 years or older, had their records age corrected also, but such records were only allowed to contribute to herdmate and not to daughter averages.

Corrections for age at calving were designed to estimate the likely production of the young or very old cows had they been mature (i.e., 5 to 9 years old) and made a production record under identical environmental conditions (Searle and Henderson, 1959). That is, corrections for age were not meant to be predictions of the future production of young cows when they ultimately reached maturity or the past production of old cows. The use of age corrections allowed the daughters of an individual sire to be compared on a within-herd basis, to an average which included all herdmates milked in the same year. This comparison with all herdmates was expected to increase the accuracy of the evaluation of a sire.

Despite offering an opportunity for increasing the accuracy of sire evaluation, the use of age corrections came into disfavour through introducing errors into progeny testing (N.Z. Dairy Board, 1970). Correcting for age led to a sire-by-herd interaction of a magnitude sufficient to bias sire evaluation (Rae, 1971). The reason why age corrections were the cause of sire-by-herd interaction requires amplification as set out below.

Kay and M'Candlish (1929) noted that in estimating productive changes associated with age, the age of an animal was used as an index of production at maturity. Searle and Henderson (1959) suggested that corrections for age may be more accurate if production levels of individual herds were taken into account in the age-correcting procedure because of herd differences in rearing and feeding, particularly as these affect the performance of young animals. They obtained support for their suggestion through significant regressions of the age corrections on the age-corrected herd averages for all age groups being corrected; the data were obtained from herds in the State of New York.

Other studies have demonstrated a relationship between the production level of the herd and the actual amount by which the production of young animals changes with age. Ward and Campbell (1938) found 2-year-old cows in low producing New Zealand herds to increase in production by a greater amount than their counterparts in high producing herds. Conversely, Castle (1955), as cited by Searle and Henderson (1959), found the intra-herd productive difference between young and mature cows in New Zealand to be greater in high producing herds than in low producing herds. In the United States, Hickman and Henderson (1955) obtained a negative correlation between the level of first-lactation production and the subsequent increase in production when a cow entered her second lactation, a finding which supports that of Ward and Campbell (1938).

Inter-herd differences in the amount by which production alters with age form the basis of an age-by-herd interaction component in age-corrected records if the correction procedure disregards these differences. Herd-level age-correction factors, developed by Searle and Henderson (1959), appeared by design to take account of this variation and thus any interaction that might exist between herd environment and the effect of age at calving on production. However, Searle (1962a) demonstrated that in terms of minimizing an age-by-herd interaction component in corrected records, the simplified herd-level factors in use in New Zealand (Searle, 1960) were only slightly superior to another correction procedure not designed to take account of the requirements of individual herds.

If an age-correcting procedure is accompanied by an age-by-herd interaction component in the corrected records, and if the daughters of the sires being evaluated are unevenly distributed amongst a series of herds, then sire-by-herd interactions are likely to be introduced or, if already present, magnified. Accordingly, R. H. Miller, McDaniel and Dickinson (1970) suggested age-correction factors be calculated on a within-herd basis. Whilst this may overcome the problem of an age-by-herd interaction component in the corrected records, from a practical standpoint the suggestion may be unrealistic due to the amount of computation involved.

The problem of an interaction between environment and the amount by which production changes with age can be extrapolated to the regional level. Studies by the N.Z. Dairy Board (1961) and the N.Z. Jersey Cattle Breeders' Association (N.Z. Dairy Board, 1961) demonstrated that increases in production of young animals with age were not uniform in all districts of New Zealand. The districts, each represented by a Herd Improvement Association, differ in features such as climate, soil type and management policies and therefore, considering the discussion so far, regional variation in the productive change due to age is to be expected. Accordingly, age-correction estimates were computed separately for each Herd Improvement Association (N.Z. Dairy Board, 1961). In the United States, R.H. Miller (1964) detected regional variation in age-correction estimates. Similar findings were obtained by McDaniel and Corley (1966) who, along with P.D. Miller, Lentz and Henderson (1970) concluded that biases in sire evaluation would result if regional variation in age-correction requirements was not taken into account. The use of age-correction factors specific for a Herd Improvement Association implied that the required age corrections for each herd within an Association were the same, but the validity of that implication has been questioned above.

Accurate age correcting of lactation records is confronted by other difficulties. For example, the unbiased estimation of age-correction factors is complicated by the practice of culling on the basis of previous productive performance (Lush and Shrode, 1950). Henderson (1949) introduced a method, referred to as "maximum likelihood", for estimating fixed effects (such as age effects) in situations where selection has been practiced and, more especially, where productive performance is incompletely repeatable. Where large amounts of data are involved, Henderson's method is difficult to apply from a computational standpoint. Despite this drawback, R.H. Miller, Harvey, Tabler, McDaniel and Corley (1966), P.D. Miller and Henderson (1968), P.D. Miller <u>et al</u>. (1970) and Wickham (1972) have successfully applied

Henderson's method for estimating age effects using a variety of linear models. Whether maximum likelihood estimates of age effects, which are really best linear unbiased estimates (Henderson, 1973), are free from the problem of age-by-herd interactions, has yet to be determined. In the case of R.H. Miller et al. (1966), for example, age-by-herd interactions were simply assumed to be non-existent.

The use of age corrections can be avoided by a within age (or contemporary) comparison of the average production of a sire's daughters with their herdmate average. The contemporary comparison is usually confined to first-lactation records to avoid biases resulting from practices such as differential culling of the daughters of different sires. Recently, McDaniel, Norman and Dickinson (1973) studied the merits of the contemporary comparison as opposed to comparisons involving all herdmates for sire evaluation. They concluded that the main disadvantages of the use of the contemporary comparison, especially where the average herd size is small, were:

(i) A loss of useful information in some herds due to a lack of paternally unrelated contemporaries to which the daughters can be compared, or

(ii) less precision in sire evaluation since the contemporary comparison involves fewer animals than if all herdmates are included in the comparison.

In most New Zealand herds, the number of daughters and contemporaries is sufficient to allow the use of the contemporary comparison in sire evaluation (N.Z. Dairy Board, 1970). The average size of New Zealand herds is larger than that of most countries and as a result, the above-mentioned disadvantages of the contemporary comparison are unlikely to be as important.

The use of age corrections to conduct sire evaluations in New Zealand is generally unnecessary and therefore, the biases associated with their use can be avoided. The practice was discontinued in the 1969-70 dairying season.

## B. Intra-Sire Regression of Daughter Performance on Herdmate Performance

Effectively, the intra-sire regression of daughter production on herdmate production is an adjustment for the non-random usage of the sires being evaluated amongst herds of differing production levels (Henderson, Carter and Godfrey, 1954). Thus, a possible source of bias in sire evaluation resulting from genetic differences between herds should be overcome by the use of an intra-sire regression of daughter average on herdmate average.

The value of the intra-sire regression of daughter performance on herdmate performance is expected to vary in accordance with differing circumstances. For instance, if differences in herd production are entirely genetic in origin, the regression has a limiting value of 0.5, whereas, if herd differences are entirely environmental, the limiting value is 1.0 (Robertson and Rendel, 1954). According to Pirchner and Lush (1959), the continued use of A.I. will eventually eliminate all the genetic variation among herds. That the regression may change in value according to how the herdmate average is defined was suggested by R.H. Miller (1970). Support for this suggestion came from an empirical demonstration (Bereskin, 1963, [illegible] R.H. Miller, 1970), of a decline in the regression when records of close relatives of the daughters were included in the herdmate average.

An examination by Pirchner and Lush (1959) of the theoretical expectation of covariance and variance terms of the intra-sire regression coefficient of daughter production on their herdmate averages, where the herdmate average excluded only the records of the daughters and their half sisters, showed that the value of the regression estimate is dependent upon herd size. If, however, the daughter performance is regressed on the adjusted herdmate average as defined by Heidhues, Van Vleck and Henderson (1961), the regression does not depend on herd size (Mao, Henderson and P.D. Miller, 1972).

The New Zealand estimate of the intra-sire regression of daughter performance on herdmate performance for milkfat yield was 0.9 and implied that, on average, 20 percent of the between-herd difference in production was genetic in origin (Rae, 1971). The results from a number of studies support the use of a general estimate of the regression of 0.9. Searle (1961a), studying age-corrected milkfat records produced by both A.I. and naturally-bred cows in the New Zealand dairying seasons of 1956-57 (2,901 daughters, 51 sires) and 1957-58 (4,195 daughters, 62 sires), estimated the regression to be 0.94 and 0.90, respectively. In a further study, using records produced only by A.I.-bred cows (10,550) in the 1959-60 season, Searle (1963) estimated the regression to be 0.83. Searle suggested the smaller estimate of the latter study may have been due to the fact that the records analysed were from cows sired by A.I. exclusively. The basis of this suggestion may have been due to the estimate of 0.6 of Henderson et al. (1954) when the milkfat records of only A.I.-bred cows (7,850 daughters, 128 sires) were analyzed. The estimate in the study of Henderson et al. (1954) was based on the regression of first-lactation performance on a contemporary herdmate average. Later, using a larger body of data (10,292 A.I.-bred daughters, 595 sires), Henderson and Carter (1957) obtained an estimate for milkfat yield of 0.911 and admitted that the former value of 0.6 was ". . . surprisingly low". Henderson and Carter did not explicitly define the herdmate average used in their study, but since they concluded ". . . the accuracy of progeny tests can be improved materially by expressing each record as a deviation from the average of all other cows freshening in the same herd, year and season.", one presumes that such an average was used in their calculations. If so, the difference between the estimate of Henderson et al. (1954) (0.6) and that of Henderson and Carter (1957) (0.911), where a contemporary herdmate and herdmate average were used respectively, adds further support to the suggestion of R.H. Miller (1970) mentioned earlier, that the value of the regression may depend on how the herdmate average is defined. An analysis by Robertson and Rendel (1954) of milk production records made in Great Britain gave a regression estimate of 0.95. In addition, breed differences were revealed through the estimates of 0.9, 0.91 and 1.08 for the Shorthorn, Guernsey and Friesian breeds, respectively.

Estimates of the regression for milk yield and milkfat yield have usually been similar in size. As far as fat percentage is concerned, Robertson and Rendel (1954) obtained a lower regression estimate of 0.5.

In proposing a reason for the low estimate of the regression in the case of fat percentage, Robertson and Rendel stated: "Probably a fair proportion of the variation between herds is therefore genetic in origin. This is in line with evidence that the effect of increased feeding in commercial herds is an increase in yield with little change in fat content."

In a study of 418,139 mature-equivalent United States records R.H. Miller (1970) obtained breed, age at calving, trait (milk yield, milkfat yield and fat percentage) and regional differences in the estimate of the regression of daughter performance on herdmate performance. In line with Robertson and Rendel, R.H. Miller obtained an average estimate for fat percentage (0.7) which was lower than that for milk and milkfat yield (0.9). An average value of 0.9 for milk yield, whilst being suitable for the Holstein breed, was found by R.H. Miller to be too high for the Ayrshire, Jersey and Brown Swiss breeds but too low for the Guernsey breed. As far as age at calving is concerned, Miller drew attention to a " . . . striking . . . " increase in the regression from first to second lactation. He concluded that the use of a general regression estimate of 0.9, the representative value for milk or milkfat yield in his study, may introduce biases into sire evaluation and that the accuracy of sire evaluations is likely to increase if the intra-sire regression of daughter performance on herdmate performance is made more appropriate to the conditions of each sire's evaluation. Earlier, Henderson and Carter (1957) had observed breed, year, and seasonal variation in the regression estimates, but noted that the differences were too small for significance.

The claim by Pirchner and Lush (1959) that continued use of A.I would eventually eliminate all genetic variation among herds was investigated by Mao <u>et al</u>. (1972). Mao <u>et al</u> (1972) analysed approximately 900,000 milk records of Holstein cows sired by A.I. obtained over the period 1956 to 1968. Their analysis produced a pooled estimate of 0.933 (± 0.002) for the intra-sire regression of daughter performance on herdmate performance with no significant trend in the yearly estimates. Thus, this analyses did not support the claim made by Pirchner and Lush (1959) that the genetic variance between herds would be progressively reduced. Nevertheless from a linear regression of relative genetic variation among herds on years (weighted by inverse of variance within years), assuming no covariance between herd genetic and non-genetic variates, Mao et al. were able to demonstrate that, in fact, the additive genetic variation relative to total variation among herds had been decreasing by 0.6 percent a year since 1959 in the Holstein A.I. population in North-eastern United States.

The use of a single overall estimate of the intra-sire regression assumes that the regressions are linear and homogeneous for all sire groups. From an analysis of first-lactation records made by 44,975 A.I.-bred cows, Van Vleck (1963) found the intra-sire regression to be statistically heterogeneous over sire groups and that curvilinear regression accounted for a statistically significant reduction in variation above that accounted for by linear regression. Van Vleck concluded that from a practical standpoint, unless the herdmate average deviated widely from the breed average, the use of a common linear intra-sire regression of 0.88 was satisfactory for adjusting average daughter yield for herdmate level.

Notwithstanding the support these estimates give to the use of an intra-sire regression of 0.9, in 1969 the N.Z. Dairy Board discontinued the use of an intra-sire regression to adjust daughter averages for herdmate average (N.Z. Dairy Board, 1970). A number of factors may have contributed to this decision:

(i) Apart from the analyses of Searle (1961a, 1963), which involved a comparatively small body of data, all the estimates cited were from analyses of data obtained in other countries, but the system of dairyfarming in New Zealand is sufficiently different to that of other countries to enable one to question the applicability to New Zealand conditions of regression estimates derived from overseas data.

(ii) The extensive studies of R.H. Miller (1970) and Mao et al. (1972) which confirmed the estimate of 0.9 were unpublished when the use in New Zealand of an intra-sire regression approach to adjust daughter averages for the genetic level of their herdmates was discontinued.

(iii) A study of New Zealand data by Brumby (1961) further suggested that the regression should be unity since, in the case of milkfat yield, genetic differences between herds were non-existent.

Notwithstanding the above points, the results of a study by Evans (1969) probably played the major role in bringing about the discontinuation of the use of the intra-sire regression approach to adjusting for herdmate level in New Zealand. Using the reference bull(b) approach, Evans demonstrated that genetic differences between herds were not uniform and that they were dependent upon the extent to which A.I. had been used in the herds being compared. In addition, he demonstrated that the genetic level of the herdmates could be accurately predicted from a knowledge of the genetic rating of their own sires. (This is considered in more detail in Chapter 2, Section III.)

## C. Intra-Sire Regression of Future Daughter Production on Estimated True Daughter Level

The factor F (see p. 7 ), the coefficient of the regression of a sire's future daughter production on his estimated true daughter average,was separately obtained for each sire from the expression n/(n + 15) where n was the number of daughters contributing to a sire's daughter average (Searle, 1959, 1964a). Factor F can also be interpreted as the regression of one-half the sire's additive genetic value on the difference from expectancy of his daughters. However, it is incorrect to interpret the regression the other way round as in Searle (1959), that is, " . . . the regression of the "difference from expectancy" on one-half the sire's additive genetic value." The formula for obtaining the regression of one-half the sire's additive genetic value on the difference from expectancy of his daughters is:

---

(b) The reference bull technique requires that the genetic merit of a bull be separately estimated in the two herds (herd groups) being compared. Since the genetic merit of a bull is constant, the difference between the two estimates of his genetic merit represents one-half the genetic difference between the herds.

$$\frac{1/4\, h^2 n}{1+(n-1)t}$$

where: $h^2$ = heritability

$t$ = the intra-class correlation amongst the phenotypes of paternal half sisters consisting of two components: $1/4\, h^2 + c^2$.

This formula is valid only if first-lactations of the sire's daughters are used to obtain the difference from expectancy. The additive genetic variance which paternal half sisters have in common, is $1/4\, h^2$, whereas, $c^2$ represents the extra correlation due to factors other than the sire which a set of half sisters may have in common, but which differ for each set of half sisters.

Since $c^2$ is in the denominator of the formula for the regression of one-half a sire's additive genetic value on his daughters' difference from expectancy, the accuracy of the prediction of the additive genetic value is increased if $c^2$ is decreased. If in fact $c^2 = 0$, and if $h^2 = 0.25$, the above formula simply reduces to $n/(n + 15)$. A way of achieving a reduction in the size of $c^2$, to the point where it may be considered to be zero, is through arranging for each daughter of the sire being evaluated to be recorded in a different herd.

That the extra correlation amongst the records of paternal half sisters, $c^2$, can be important in affecting the accuracy of sire evaluation has been demonstrated. Commonly, $c^2$ is broadly stated as being determined by "environmental" factors. However, Bereskin and Lush (1965) drew attention to other determinants of $c^2$, these are:

(i) Correlations amongst the additive genetic values of the mates of a sire.

(ii) Correlations between the additive genetic value of the sire and the average additive genetic value of the cows to which he was mated.

(iii) Correlations involving both genetic and environmental effects.

In a comparison of three methods of distributing the daughters of a sire amongst a series of herds, Bereskin and Lush (1965) studied the effect of $c^2$ on the repeatability of sire proofs. The method of interest here was that where the herds were chosen at random, and where each herd often had only one daughter of the sire being evaluated. This system of distributing the daughters was regarded by Bereskin and Lush to be " . . . very similar to that prevailing under ideal A.I. conditions." Despite the fact that their data approached the ideal situation of one daughter of the sire per herd, Bereskin and Lush demonstrated that the use of production deviations (obtained from comparing daughter and herdmate averages) did not completely eliminate residual herd-year-season interaction variance from their data and that, in turn, this variance resulted in an extraneous correlation ($c^2$) of magnitude sufficient to bias sire evaluation. Accordingly, Bereskin and Lush noted the advisability of having only a single daughter of the sire in each herd.

In a study of first-lactation records, Van Vleck (1966b) estimated the environmental correlation amongst the records of paternal half sisters recorded in the same herd to be 0.08. Since the fraction of daughters of an A.I. sire in the same herd relative to the fraction in different herds is small, Van Vleck concluded that the correlation is unlikely to be important in the evaluation of A.I. sires. This conclusion was supported by his earlier demonstration (Van Vleck, 1966a) that the correlation may not be important in A.I. sire evaluation in the United States. However, in evaluating sires by natural service, the distribution of daughters is a direct opposite of that for A.I. sires and, in both studies, Van Vleck concluded that account should be taken of the correlation $c^2$ in the evaluation of natural service sires.

The lack of concordance between natural service and A.I. sire evaluations, observed in New Zealand and elsewhere, is attributable in part to the effect of the correlation $c^2$ (Van Vleck, 1966a). Arora and Freeman (1970) obtained a common value for the correlation $c^2$ of 0.11 for A.I. and non-A.I. populations and concluded that the correlation should be taken into account in sire evaluation.

If each daughter of the sire being evaluated is located in a different herd, and if sires are used at random in a series of herds, then it is reasonable to assume that the correlation $c^2$ is zero. As far as the New Zealand method of evaluation is concerned, $c^2$ can be assumed to be zero since the foregoing requirements seem to be met through the N.Z. Dairy Board's sire proving scheme as described in the following sub-section.

In the derivation of the regression coefficient n/(n+15), a requirement which needs to be met is that only first lactations of the sire's daughters be used to obtain their difference from expectancy. Searle (1964a), using the method of Henderson, Kempthorne, Searle and von Krosigk (1959), demonstrated that culling, based on previous lactational performance, was responsible for biases in sire evaluation if second or succeeding daughter lactations were used for calculating differences from expectancy. However, Searle (1959, 1964a) developed an alternative expression which allowed the bias to be avoided if second or later lactations were included in sire averages. Even though the bias resulting from culling can be overcome, the use of second or later lactations is inadvisable. For instance, the use of second or later lactations means that records have to be age corrected which, as already indicated, can lead to other difficulties.

## D. New Zealand Dairy Board's Sire Proving Scheme

An integral part of the New Zealand method of sire evaluation is the N.Z. Dairy Board's sire proving scheme, introduced in 1961 (N.Z. Dairy Board, 1961). Under this scheme, inseminations from young unproven bulls are distributed amongst a large number of sire proving herds. Farmers participating in the scheme are paid by the Dairy Board not only for using semen of unproven bulls, but also for guaranteeing to record all daughters that come into milk three years later. Prior to 1961, inseminations from unproven sires were distributed amongst all herds using A.I. However, under these circumstances, only five tested daughters were obtained for every 100 inseminations although, on average, 100 inseminations produced 21 milking 2-year old heifers. The discrepancy between the number tested and the number of daughters actually in milk, resulted from the fact that not all herds using A.I. were herd recording.

It was demonstrated (N.Z. Dairy Board, 1961) that testing all daughters of an unproven bull reduced the number of inseminations required to produce a sufficient number of daughters on whose productive performance the bull's evaluation was based. Searle (1962b) gave the reduction in the number of inseminations to be approximately 700 per unproven bull. A reduction in the number of inseminations from each unproven bull allowed more bulls to be evaluated, but because the number of replacement bulls required for the proven bull team remained the same, a higher selection differential could be achieved. The increased selection differential, coupled with an overall reduction in the use of unproven semen, was expected to bring about an increase in the rate of genetic gain in the A.I. population (N.Z. Dairy Board, 1961).

## III. SIRE EVALUATION IN NEW ZEALAND SINCE 1969

The N.Z. Dairy Board continued to use the method of sire evaluation as outlined on page 7 for a number of years before introducing major changes in 1970 to the method (N.Z. Dairy Board, 1970). There were several aspects relating to the changes that were made and these will now be considered.

Prior to 1970, the herdmate averages to which the daughters of the sire being evaluated were compared were calculated using only the production records of naturally-bred herdmates (Searle, 1964a; N.Z. Dairy Board, 1970). In contrast to A.I.-sired cows, naturally-bred cows were assumed to have been sired by bulls whose average genetic merit equalled that of the population. Through including the records of only naturally-bred cows in herdmate averages (thereby excluding the records of all A.I.-sired cows), it was assumed that a sire's estimated genetic merit (expressed relative to the population average) was unbiased by the genetic superiority of any A.I.-sired herdmates (N.Z. Dairy Board, 1970). That New Zealand A.I.-sired cows are genetically different to naturally-bred (assumed to be non-A.I.-sired) cows has been demonstrated by Evans (1969) and by the Farm Production Division of the N.Z. Dairy Board (1974). The results of the latter study are given in Table 2.2.

Farmers making use of A.I. in their herds often retain sons of A.I. sires for use in natural mating. Thus, although some cows are identified as being naturally bred, in fact they have A.I.-bred ancestors and, for reasons already discussed, these cows are unlikely to be genetically average. Since A.I. sires are not uniformly used in a series of herds, a bias in sire evaluation may occur due to variation in the genetic averages of the herdmates to which the daughters of the different sires are compared.

Studies in the United States by Norman, McDaniel and Dickinson (1971) and Henderson and co-workers (unpublished) have demonstrated that variation in the genetic levels of herdmates does occur in practice. The studies of Van Vleck, Henderson and Carter (1962), R.H. Miller and Corley (1965), Bereskin and Freeman (1965), Hillers and Freeman (1966), and Bodoh and Shook (1972) indicate that, although the intensity with which the mates of each sire have been selected may be significantly different from sire to sire, this is an unimportant source of error in sire evaluation. Furthermore, Rønningen (1971) showed that the production deviation of a sire's mates, together with the heritability of the trait, must be very high before the records of the mates need to be considered. Thus, variation in the genetic level of herdmates is more likely to be due to a differential genetic merit of the sires and <u>not</u> the dams of the herdmates. Accordingly, Norman <u>et al</u>. (1971) suggested herdmate averages should be adjusted by some function of the genetic merit of the herdmate's sires.

From the point of view of adjusting herdmate averages according to the genetic merit of herdmate's sires, a study of importance to sire evaluation in New Zealand was that of Evans (1969). On the basis of the ratings of proven A.I. Jersey sires, Evans calculated an expected rating for each of 223 sons of these proven sires. The son's observed

Table 2.2 : New Zealand Dairy Board's within-herd production comparison of identified A.I.-sired cows with other (non-A.I.) cows in the 1972-73 dairying season.

| Age class | Production difference (in favour of A.I.-sired cows) | |
|---|---|---|
| | Milk (lb) | Milkfat (lb) |
| 2 years | 458 | 23 |
| 3 years | 333 | 21 |
| 4 years | 399 | 21 |
| 5 years and over | 297 | 15 |

milkfat ratings, obtained from their own progeny tests, were then regressed on their expected ratings. A regression coefficient of 0.96 was obtained which, when compared to an expected value of 1.0, indicated that the prediction of the genetic merit of sons of proven A.I. sires was on average 96 percent efficient. On the other hand Evans found the production information of the dam's of the sons predicted the son's genetic merit with much less precision, being on average only 50 percent efficient.

The findings made by Evans (1969) of a strong relationship between the estimated genetic merit of proven sires and their sons is important since, as noted by Robertson (1960), the success of an A.I. scheme relies on the identification of young bulls which are above average in genetic merit. Robertson (1960) introduced the idea of regressing the progeny test result of a son on that of his sire to assess the accuracy of a sire evaluation procedure in identifying sons of high genetic merit. In the case of milk yield, Robertson (1960) demonstrated that the contemporary comparison method of sire evaluation (MacArthur, 1954; Robertson, Stewart and Ashton, 1956) was only 70 percent efficient in identifying sons of above average genetic merit. This finding was confirmed by the study of Ødegård and Robertson (1968). In contrast, studies by McDaniel, 1969; Dickinson, McDaniel, R.H. Miller and Lytton, 1969; and Ødegård, 1968, have shown that the observed regression of son's progeny test on that of their sires, approximated the expected regression for milk yield, indicating that the sire evaluation procedures concerned were operating very close to their theoretically expected precision. Thus, the ratings of proven sires appear, on average, to be a reliable basis for the prediction of the genetic merit of their progeny. Nevertheless, Searle (1964b) has shown that in selecting sons of proven sires a knowledge of the sire's genetic merit should not be substituted for a progeny test of the sons themselves. Searle (1964b) demonstrated that for a heritability of 0.25 for a trait, the most accurate evaluation of a sire's genetic merit provides no more information about a son's true genetic merit than does a progeny test based on five daughters of the son.

The demonstration by Evans (1969) of the low utility of a dam's milkfat production records as a means for predicting a son's genetic merit is in accord with the results of studies elsewhere. Using the selection index approach for predicting the genetic merit for milk yield of the daughters of a young bull, on the basis of information at various points in his pedigree, Van Vleck (1969) demonstrated that exclusion of dams' records brought an unimportant change in the efficiency of the index. Similarly, Van Vleck and Carter (1972) and Lindström and Maijala (1972) showed a low relationship between dams' milk yield information and the genetic merit of their sons. These results do not mean that information on the dams is of no use for predicting the genetic merit of their sons in order to distinguish those that should be progeny tested for possible use in A.I. Rather, they imply that since dams of young bulls are usually a selected group, differences among them are of little consequence in predicting the genetic merit of their sons (Van Vleck and Carter, 1972). Moreover, Van Vleck (1969) found the most efficient selection index to be one that included the records of the dams, i.e., the proofs of the sire and maternal grandsire of the young bull and the production performance of his dam. Van Vleck and Carter (1972) found this to be " . . . a very effective method of gathering a superior group of young bulls for further sampling in A.I. . . . "

The method of sire evaluation introduced by the N.Z. Dairy Board in 1970 utilizes ancestral information and, in particular, capitalizes on the reliability of the estimated genetic merit of a proven sire in predicting the genetic merit of his progeny. The outline of the method (N.Z. Dairy Board, 1970) is as follows:

$$\text{Sire's rating} \quad = \quad \text{Sire's } \frac{\text{E.G.V.}}{2} + \text{R (Difference)}$$

where: E.G.V. = The expected genetic value of the sire, obtained from the estimated (expected) genetic value of his ancestors, notably the ratings of his sire and maternal grandsire plus the production record of his dam. When ancestral information is unavailable, i.e., when no ancestors are A.I.-bred, the sire's E.G.V. = 0.

Difference = (Daughter average - Herdmate average) - (Daughters' E.G.V. - Herdmates' E.G.V.)

= The within-herd, within-age average production deviation of the sire's daughters from their herdmates, adjusted by the expected difference between the genetic averages of the two groups.

R = The regression coefficient of adjust the deviation of a sire's actual rating from his expected rating.

Calculation of the herdmates' E.G.V. estimates the genetic level of the herd and thus the need for an intra-sire regression of daughter performance on herdmates' performance to take account of genetic differences between herds is overcome. The use of the herdmates' E.G.V. in calculating ratings should avoid the bias in sire evaluation due to unequal genetic merit of the herdmates. If no reliable ancestral information is available (i.e., no A.I. ancestry amongst the herdmates) the genetic level of the herd (Herdmates' E.G.V.) is considered to be zero.

The use of the coefficient n/(n + 15) (factor F) for the regression of the genetic merit of an unproven bull on his true daughter average is inappropriate when the bulls being evaluated are initially selected on the basis of ancestry, i.e., not chosen from the population at random (Rae, 1971). The value for R is obtained, according to the circumstances, from the use of a selection index to combine information from different sources (e.g., sire's rating of the unproven bull, unproven bull's true daughter average, production records of the unproven bull's dam) to estimate the genetic merit of the unproven bull.

The accuracy of the above method of evaluation, particularly insofar as the ranking of older sires in comparison to sires in younger age groups is concerned, is a major aspect of this thesis.

# IV. GENETIC TREND

## A. Genetic Trend as a Bias in Sire Evaluation

Van Vleck and Henderson (1961a) estimated the increase in the genetic merit of the Holstein population of New York State in the period 1951 to 1959 to be approximately 0.5 percent of the population production mean per year. Subsequently, Van Vleck and Henderson (1961b) found that the ranking of 38 sires with daughters having lactated in that period (1951 to 1959) was only slightly altered by the adjustment of each sire's estimated genetic merit for the effects of genetic trend. However, by present-day standards a genetic trend of 0.5 percent of the population mean per year is generally regarded as being low. Harville and Henderson (1967) estimated that the genetic merit of the Holstein population in New York State increased by 0.7 percent of the mean per year in the period 1956 to 1962 and demonstrated that the correlation between the estimated genetic merit of sires before and after adjustment for genetic trend effects was almost unity, a result which supports the findings of Van Vleck and Henderson (1961b). In addition, however, Harville and Henderson demonstrated that the ranking of certain categories of sires was significantly altered by adjustment for the effects of genetic trend. According to Henderson (1973), the level of genetic trend in the cow population of New York State by 1966 was responsible for sufficient bias in the herdmate comparison method of sire evaluation (Heidhues et al., 1961) to justify the introduction of methods unbiased by genetic trend. The way to overcome the bias is by judicious grouping of sires by making direct comparisons between sires as opposed to comparisons with herdmate or contemporary averages (Lentz, P. D. Miller and Henderson, 1969). A method which utilizes the direct comparison approach is studied in this thesis.

## B. Methods for Estimating Genetic Trend

The major impetus behind the development of methods to estimate genetic trend has come from the desire to determine the success of selection programmes. However, the estimation of genetic trend, unconfounded with environmental effects and other biases, is not straightforward. Basically, three approaches to this estimation problem have been used for dairy cattle populations. The approaches are:

(1) Prediction from theory. The average genetic gain per year in a population can be predicted from knowledge of:

(i) The average genetic superiority of the animals selected to be parents over the group whence they came and;

(ii) the average generation interval in the selection programme (Dickerson and Hazel, 1944).

The genetic superiority and the generation interval may be different for males as opposed to females so that in applying the method of Dickerson and Hazel to a dairy cattle population, Rendel and Robertson (1950) obtained these parameters separately for each of the four pathways by which genes are transmitted to a succeeding generation. That is:

$$G = \frac{I_{\text{Sire to son}} + I_{\text{Sire to daughter}} + I_{\text{Dam to son}} + I_{\text{Dam to daughter}}}{L_{\text{Sire to son}} + L_{\text{Sire to daughter}} + L_{\text{Dam to son}} + L_{\text{Dam to daughter}}}$$

where:

$L$ = Generation interval.

$I$ = Genetic superiority of the animals selected to be parents over the average of their contemporaries.

$I$, which is also the genetic selection differential, is calculated for each pathway using the formula:

$$I = \bar{i}\, h\, \sigma_g$$

where:

$\bar{i}$ = Phenotypic superiority of the selected animals expressed in standard units.

h = Square-root of heritability.

$\sigma g$ = Additive genetic standard deviation.

The prediction approach for estimating genetic gain has often been used to provide a basis on which the efficiences of alternative selection programmes can be compared since the criterion of success in a selection programme is normally the genetic gain the programme is likely to produce. Robertson and Rendel (1950) demonstrated that the efficient use of progeny testing in association with A.I. should, theoretically, be capable of producing rates of genetic gain per year of 1.69 and 2.05 percent of the population mean for milk yield in populations of 2,000 and 10,000 cows, respectively. Specht and McGilliard (1960) found that for an A.I. population of 10,000 cows the rate of genetic gain obtained from progeny testing varied between 1.7 and 2.3 percent of the population mean per year depending upon the number of young sires sampled each year and the proportion of these sires that were ultimately selected for routine use in A.I. Based on a population size of 60,000 cows, Legates (1971) concluded that the judicious use of progeny testing in association with A.I. should produce a genetic gain per year of 1.8 percent of the population mean. In a study of the optimum utilization of young bulls in an A.I. scheme involving 66,000 cows, Skjervold (1963) obtained a highest value of 1.8 percent of the mean per year for the genetic gain in milk production which is the same as the estimate obtained by Legates (1971).

A number of research workers have employed the technique of Robertson and Rendel (1950) to predict the amount of genetic gain achieved. Rendel and Robertson (1950) assessed the genetic gain per year in a closed herd not making use of progeny testing to be 0.7 percent of the herd average for milk yield and estimated the maximum possible genetic gain in such a herd to be 1.0 percent of the average per year. For a closed herd of Red Sindhi cattle, Stonaker (1953) predicted the genetic gain due to selection to be 0.3 percent of the herd's milkfat average per year. Based on data from a closed herd of Hariana cows, Acharya and Lush (1968) obtained an unusually high predicted genetic gain of 2.5 percent of the herd average per year.

This was attributed to the high values for the heritability and repeatability which Acharya and Lush calculated from the data.

From an analysis of first-lactation records of the same herd, using a method unbiased by inaccurate estimates of heritability, repeatability or of age corrections (Smith, 1962), Acharya and Lush (1968) obtained an estimate of the actual genetic gain of 1.5 percent of the herd average per year. Even so, considering the estimate of Robertson and Rendel (1950) of a maximum possible genetic gain in a closed herd of 1.0 percent of the herd average per year, the latter estimate of Acharya and Lush of 1.5 percent is again high.

The prediction method has also been used to determine the genetic gain being achieved in selection programmes making use of progeny testing and A.I. Syrstad (1966) estimated the expected genetic gain per year in the Norwegian Red breed in the period 1957 to 1964 to be approximately 1.0 percent of the average milk yield. The expected genetic gain currently being achieved by the N.Z. Dairy Board's progeny testing and A.I. scheme has been calculated by Jackson (1974). After adjustment for a rise in the level of inbreeding, Jackson obtained an expected genetic gain of 6.8 lb. of milkfat per cow per year which, assuming a mature-equivalent average production of 350 lb. of milkfat, represents a genetic gain of 1.94 percent of the population average per year. The structure of the N.Z. Dairy Board's breeding scheme was outlined by Evans (1969) and shown to involve approximately 1 million cows. Therefore, bearing in mind the studies of Robertson and Rendel (1950); Specht and McGilliard (1960); Skjervold (1963); and Legates (1971), an expected genetic gain per year for the New Zealand scheme of 1.94 percent of the mean (Jackson, 1974) is not unreasonable.

The major limitation of the prediction method for estimating genetic gain is that it is based on a simple model which assumes the genetic and phenotypic parameters to be known, all genetic variance to be additive, and all selection is applied to the single trait being considered. The rates of genetic gain predicted by this method have seldom been achieved in practice, a point which is detailed in the ensuing discussion.

(2) Estimation using a control population. An estimate of the genetic gain achieved through selection can be obtained from a comparison of the phenotypic performance of the selected population with that of a randomly-bred control population maintained in the same environment.

Legates and Myers (1966) and Legates (1971) have described the use of this technique to determine the genetic progress, from 1955 onwards, in a selected group of cows in the North Carolina State University herd. The estimates of the annual genetic gain in milk yield in the two studies were 176 and 245 lb., respectively or, alternatively, between 1.0 and 1.5 percent of the herd average per year (Legates, 1971). On the basis of the technique of Smith (1962) for estimating genetic change, Legates (1971) checked the reliability of his estimate (245 lb per year) by calculating the genetic gain per year in six state-owned herds which had used the same sires as those used in the University's selected group. The estimated genetic gain in the state-owned herds was 232 lb. of milk per year, an estimate which agrees rather well with that obtained in the University herd where the control population approach for estimating genetic gain had been used. Hill (1972b) in his review of the work of Legates and Myers (1966) commented: "The population [control] has been maintained alongside one selected for fat yield, but has only 6 bulls per generation, with about 20 cows in lactation, so trends will be difficult to detect accurately."

The control population approach for estimating genetic gain has rarely been made use of in genetic studies of dairy cattle populations. Most probably, the major deterrent of the approach would be the cost of maintaining special groups, of reasonable size, in which no genetic gain is being made. That the control group needs to be of reasonable size is made obvious by the studies of Hill (1972a, 1972b). In a review of the general theory and design of control populations, Hill (1972a) drew attention to the possible sources of error associated with this approach for estimating genetic gain. These are: Random genetic drift in the control population; genetic trend in the control population as a result of natural selection; genotype-environment interaction (differential response of selected and control populations to environmental changes); and the error of estimation of the genotypic mean on the basis of the phenotypic mean of both populations. Hill (1972b) attempted to assess the importance of these sources of error through an extensive review of selection experiments, involving predominantly laboratory species, in which control populations had been used. It would appear that in designing control populations for estimating the genetic gain in dairy cattle selection programmes, the source of error requiring most emphasis is that arising from the estimation of genotypic means on the basis of corresponding phenotypic means. The cost of maintaining populations of sufficient size to minimize the effects of this source of error probably does not justify the use of the control population approach.

(3) Estimation in non-experimental data. A number of statistical procedures for estimating the actual genetic change achieved by the application of selection to dairy cattle populations have been proposed, largely in response to the limitations of the prediction and control population methods of estimation.

One of the first attempts to estimate the genetic change in milk production, unconfounded with environmental effects, was that of Lörtscher (1937). His procedure relied on the principle that a cow's genotype is constant and that any change in her milk production with time is due to a change in her environment. Accordingly, a comparison between the age-corrected records of a given group of cows in succeeding years should reflect the change in milk production resulting from environmental trend. Once the contribution of the environment to any change in the milk production of the population as a whole has been determined, the genetic change can then be obtained by difference. Nelson (1943) refined the technique of Lörtscher (1937) by the application of least squares to a linear model in which year and cow effects were assumed to be fixed. Henderson (1949) demonstrated that because of the combined effect of cow culling on the basis of lactational performance and the incomplete repeatability of dairy cow performance, the application of least squares where cow effects were assumed to be fixed, leads to biased estimates of year effects.

Henderson's description of the bias was: "The crucial point is that least squares and similar methods are essentially year to year comparisons of records of the survivors of each year's culling. Therefore, if the survivors of culling made above herd average records prior to the culling of certain of their mates, we should expect these survivors' records to be less in succeeding years [due to incomplete repeatability] and consequently to make it appear that the environment is becoming poorer from year to year." A statistical description of the bias was given by Henderson et al. (1959) based on arguments analagous to those used by Lush and Shrode (1950) in demonstrating the way in which culling and incomplete repeatability bias age-correction factors.

If the cow effects in the model are regarded as being random effects, application of a mixed model solution method of estimation (Henderson, 1949) produces estimates of year effects which are unbiased by culling or incomplete repeatability. This amounts to having a prior estimate of the repeatability and, like the procedures of Lörtscher (1937) and Nelson (1943), Henderson's method requires that the records be age corrected before the analysis of year effects. The major limitations of this method are its sensitivity to errors in the repeatability and age-correction estimates, and the computational effort involved when applied to large sets of data (Henderson, 1953; Henderson et al., 1959). Henderson stated that: "For each pound too much by which first-lactation records are built up, the environmental trend is biased downward by 0.22 pound per cow per year. For each 0.01 by which the repeatability actually used in the estimation procedure exceeds true repeatability, the down bias is 0.08 pounds of fat per cow per year." Kempthorne and von Krosigk (in Henderson et al., 1959) developed a method for unbiased estimation of year effects of which an initial phase was the estimation of the repeatability from the same set of data. Searle (in Henderson et al., 1959) proved that for a known repeatability value, Kempthorne and von Krosigk's method is equivalent to that originally proposed by Henderson. A premise on which both methods operate is that the data have been previously age corrected using accurate age-correction factors. Rendel and Robertson (1950) demonstrated that for any given set of records, the genetic trend and age effects are completely confounded. This gives rise to the impossible situation whereby an estimate of the genetic trend is required for the accurate estimation of age-correction factors which are in turn required to estimate the genetic trend. It has only been in recent years that methods for estimating age-correction factors unbiased by genetic trend have been developed (c f. P.D. Miller et al., 1970).

Since the application of the method of Henderson (1949) for estimating genetic trend may be prohibitive from a computational standpoint, it has been used by comparatively few research workers. Moreover, the studies in which the method has been used have involved using only few data. In a study of the records made by cows in 12 California herds in the period 1930 to 1960, Arave, Laben and Mead (1964) used Henderson's method on an intra-herd basis and obtained a pooled estimate of the genetic trend of 0.7 percent of the mean yield per year. The analysis involved a total of 9,742 lactation records but the highest number of records for any one herd, and therefore the highest number for any one application of Henderson's method, was only 1692. McDaniel, Plowman and Davis (1961) proposed that the environmental trend be estimated on the basis of knowledge of the quantitative effect of changes in components of the environment, for example, feeding levels, on milk production. In order to test their proposition, McDaniel et al. (1961) used this approach to estimate the environmental trend in a single herd (230 cows, 462 lactations) in the period 1952-1958 and compared the estimate to that obtained using Henderson's technique. Despite the fact that the two estimates of environmental trend in this example were very similar, the accuracy, and also the utility, of the method of McDaniel et al. (1961), especially in a multi-herd context involving large numbers of observations, remains questionable. The method not only presumes that all the components of the environment affecting milk production can be identified, but also that their effects are quantifiable. From a practical standpoint, this is unrealistic.

A variation of the control population approach for estimating the genetic change in a selected population was introduced by Robertson and Rendel (1954). The average within-herd production deviation (weighted by the inverse of the variance of the deviation) of A.I.-bred cows, relative to their naturally-bred (non-A.I.) herdmates, is an estimate of the genetic difference between the two populations (Robertson and Rendel, 1954). In order that the estimate of the genetic trend be unbiased, the A.I.- and non-A.I. bred animals should be uniformly managed and the average genetic merit of the dams of both groups must be equal. In a study of the records of 3,152 cows (1,423 A.I.- and 1,729 non-A.I.), Robertson and Rendel (1954) demonstrated that A.I.-bred animals and non-A.I. bred animals were not genetically different in terms of production. However, at the time of their study, A.I. had not long been operating. McDaniel and King (1974) applied the technique suggested by Robertson and Rendel (within-herd comparison) to the analysis of about 4 million lactation records produced nationwide in the United States in the years 1954 to 1969, inclusive. The results of the study showed (after pooling the published estimates across cow age and breed) that by 1969 the A.I.-bred population was genetically superior to the naturally-bred population to the extent of 262 and 10 lb. of milk and milkfat per cow, respectively. Results from similar studies in the United States (Tucker, Legates and Farthing, 1960; Van Vleck and Henderson, 1961a; and Corley, Duckwall and Heizer, 1963) corroborate the findings of McDaniel and King (1974). The Farm Production Division of the N.Z. Dairy Board has also employed the method of Robertson and Rendel to compare the genetic merit of A.I.- and non-A.I. bred cows. The results of the most recent study (N.Z. Dairy Board, 1974), involving 1,298 herds and 18,600 A.I.- and 28,864 non-A.I. cows have already been given in Table 2.2.

Although the technique of Robertson and Rendel (1954) furnishes an estimate of the genetic trend of one group relative to that of another, it does not provide a reliable indication of the overall genetic trend in the entire dairy cow population. It cannot be assumed for instance, that the genetic trend in the "control population" (i.e., naturally-bred cows) is zero since farmers may follow the practice of using the sons of certain A.I. sires in natural mating, the offspring of which are classified as being naturally bred. Moreover, this practice means that the method of Robertson and Rendel is incapable of producing an unbiased estimate of the true genetic difference between A.I.- and naturally-bred populations unless the latter can be shown to have no A.I. ancestry. If the use of their method is not restricted to the analysis of first-lactation records, differential culling of A.I.- and naturally-bred cows can be another source of bias in the comparison.

Van Vleck and Henderson (1961b) introduced a modification of the procedure of Robertson and Rendel (1954) which, although computationally more complex, produces an estimate of the genetic trend in the A.I. population which is unbiased by the genetic trend in the naturally-bred (non-A.I.) population. The method proposed by Van Vleck and Henderson (1961b) is based on the application of least squares analysis of weighted contemporary deviations obtained from first-lactation records (avoiding bias due to female culling) and furnishes an estimate of not only the genetic trend in the A.I. population (corrected for trend in the non-A.I. population), but also the genetic trend in the non-A.I. population. Van Vleck and Henderson (1961a) applied their method to the analysis of first-lactation records made in New York State in the period 1950 to 1959 and demonstrated that, in fact, genetic trend did occur in the non-A.I. population as well as in the A.I. population. Although the method of Van Vleck and Henderson produces an estimate of the genetic trend in both the A.I.- and non-A.I. populations, it does not provide an estimate of the genetic trend in the total population. For the latter to be obtained, the proportion of the dairy cow population that were either A.I.- or non-A.I. bred, would have to be taken into consideration. Like the procedure of Robertson and Rendel (1954), the method of Van Vleck and Henderson relies heavily on the assumption that the A.I.- and non-A.I. bred cows within a herd are uniformly managed and that the average genetic merit of the dams of the two groups is equal.

A number of research workers have employed a repeat mating design or variation thereof to estimate the overall genetic trend actually achieved by a history of selection in a population. Effectively, the repeat mating method, first proposed for poultry populations by Goodwin, Dickerson and Lamoreux, (1955) and elaborated further by Dickerson (1960), generates a genetically constant control population through the replication of equivalent genotypes in successive years. Thus, the genetic trend can be estimated using the same principles as the control population approach already described. The basis on which repeat mating methods depend is that the genotype of an animal is constant over time and therefore, the progeny produced by repeated matings of two parents in different years have the same expected genotypic mean. The method initially proposed by Goodwin et al. (1955) relied on the assumption of no age change in parental influence on progeny performance such as maternal age effects. Dickerson (1960) gave a modification of the repeat mating method of Goodwin et al. (1955) which allows the genetic trend to be estimated free of maternal age effects. The control population approach for estimating genetic trend, in which it is necessary to assume no genetic change due to natural selection or random genetic drift in the control population was described earlier. In contrast, this assumption is unnecessary when estimating genetic trend on the basis of a control population which has been generated by a repeat mating method (Dickerson, 1960).

A study of the usefulness of the first-lactation records of full and half sisters for estimating the genetic trend in a dairy cattle population by Burnside and Legates (1967) was, in effect, an example of the application of the principles of the repeat mating method to a situation in which matings were not deliberately planned. From a total of 34,380 first-lactation records made in 335 herds in the period 1953 to 1961, Burnside and Legates (1967) generated three categories of data. These were: The records of half sisters; the records of full sisters; and the records of all animals. The half-sister data included the lactations of daughters of only those sires which were used in at least two years; in the case of the full-sister data this requirement was automatically satisfied. Burnside and Legates (1967) obtained least squares estimates of year effects from each set of data using a model in which all effects, including sire effects, were assumed to be fixed. Their justification for a fixed model was: ". . . since interest centered on the estimation of genetic trend in the particular group of herds studied." Regressions of weighted estimates of year effects on time produced estimates of: The environmental trend (full-sister data); the environmental trend plus one-half the genetic trend (half-sister data); and the environmental trend plus the genetic trend (complete data). These regressions give rise to two estimates of genetic trend and in the data of Burnside and Legates,in the case of milk production, these were: 0.75 percent of the mean per year (obtained from difference between the complete and the full-sister regressions); and 0.92 percent of the mean per year (obtained from twice the difference between the complete and the half-sister regressions). Estimates of the full-sister regression and therefore, estimates of the environmental trend, are biased if the production performance of a first (younger) full sister determines the existence of more full sisters with completed lactations.

The application of the half-sister method of Burnside and Legates (1967) to the analysis of 625 first-lactation records made in a Jersey herd in the period 1931 to 1968 by Palmer, Wilcox, Martin, Verde and Barrantes (1972),produced estimates of genetic trend of 1.3 percent and 1.0 percent of the mean production for milk and milkfat, respectively. Similarly, Verde, Wilcox, Martin and Reaves (1972) used the technique to estimate the genetic trend in 181 Florida herds in which A.I. had been used. A total of 4,779 first-lactation records were analyzed and the estimates of genetic trend in milk and milkfat production (as a percentage of the mean) varied according to breed: Holsteins 0.69 and -0.38; Jersey 0.69 and 0.82; and Guernseys 2.83 and 1.9, respectively. The validity of assuming a fixed effects model when applying the method of Burnside and Legates, particularly with respect to sire effects, is questionable. This is especially so in the case of the study of Verde et al. (1972) in which the records of A.I.-bred animals were analyzed.

In artificially-bred dairy cattle populations, full sisters are rare so that estimating the genetic trend using the full-sister comparison approach (Burnside and Legates, 1967) is likely to be impracticable in A.I. populations. The data analyzed by Burnside and Legates were derived mainly from registered naturally-bred animals with very few A.I.-bred animals represented. It is probable that full sisters in an A.I. population are the progeny of selected repeat matings, in which case, their lactation records are of little use in estimating the genetic trend because of the bias already mentioned.

Hickman and Freeman (1969) outlined a mating plan suitable for small populations whereby the frequency of full sisters resulting from random matings is maximized without a great loss of genetic change through having a herd structure which involves the use of a control population for monitoring genetic trend. In a commercial dairy cow population, parental mis-identification may occur regularly and is likely to effect the accuracy of estimates of genetic trend derived from a full-sister comparison to a greater extent than those obtained from the half-sister method since the latter only requires half the parents (the sires) to be identified. The half-sister method for estimating genetic trend as outlined by Burnside and Legates (1967) is a modification of a procedure suggested earlier by Smith (1962).

The estimation of the genetic trend actually achieved in a dairy cattle population in which A.I. is used is facilitated by the opportunity for generating half sisters over a number of years. It should be noted that since the genotype of a sire is constant, changes in the genetic merit of his daughters born in different years, represent one-half the genetic trend in the population. Since cattle semen (gametes) can be stored for long periods of time, it is possible to arrange for the progeny of sires representing different generations to be contemporaneously compared thereby providing an estimate of genetic trend (Dickerson, 1960). The formula given by Dickerson for estimating the annual genetic trend using such comparisons requires the age difference between the two groups of sires to be known. A variant of Dickerson's method which does not require knowledge of sire ages, in order that the annual genetic trend can be estimated, was used by Carter (1969). In a study involving only first-lactation records, Carter (1969) obtained an estimate of the genetic trend in a New York State dairy cow population from a comparison of the average production superiority of groups of half sisters sired by the same bulls at two points in time. Carter estimated the production superiority of the daughters of 19 young bulls randomly sampled in A.I. for the first time to be 276 lb. of milk. After a lay-off period of five years the bulls were brought back into service. The estimated production superiority of the daughters sired subsequent to the return of the bulls to service was found to be 15 lb. of milk. As the two groups of daughters are half sisters, the genetic change over the five-year period studied by Carter was 522 lb. of milk (2 x (276 - 15) or alternatively, about 104 lb. of milk per year.

Smith (1962) introduced yet another procedure which relies on the spread of half sisters over years for estimating genetic trend. The technique outlined by Smith is particularly suited to the estimation of genetic trend from large numbers of observations which have been derived under commercial conditions and consequently, it has been employed by many research workers. Although variants of the one principle, Smith (1962) outlined two methods for estimating genetic change. One of these, namely, the estimation of genetic trend from twice the difference between the regressions of the population and the pooled within-sire progeny performance on time, has already been described when reviewing the work of Burnside and Legates (1967). (In this latter study, least squares estimates of year effects estimated the population and progeny performances.) However, of the two methods outlined by Smith (1962), the one with the most utility as far as dairy cattle populations are concerned is that which estimates the genetic trend from: " . . . twice the pooled within-sire regression on time of the difference between the population and the individual sire means . . . . " The most suitable population to which the daughters of a dairy sire can be compared is a within-herd first-lactation contemporary average. Not only does the use of a first-lactation contemporary average minimize errors resulting from environmental fluctuations, it also avoids biases due to cow culling, or incorrect age-correction factors, or both.

Estimates of genetic trend obtained from the application of any of the methods which utilize the change in the productive performance of half sisters over time (Dickerson, 1960; Carter, 1969; and Smith, 1962) will be biased by non-random mating in the population. For instance, if older sires tend to be mated to older cows the estimates of the genetic trend are biased downwards. (In a population where there is a positive genetic trend, the genetic merit of the older cows tends to be below average and since the genotype of a given sire is fixed, the performance of his progeny from such matings will represent less than the total genetic change in the population. The extent of this bias is partially offset by the practice of culling females on the basis of earlier performance.) If the inbreeding coefficient of the population is increasing, the procedures which estimate genetic trend on the basis of half-sister performance over time, produce downwardly biased estimates (Hill, 1972a).

In the outline of his technique, Smith (1962) drew attention to the potential bias resulting from the selection of the sires used in a number of years on the basis of the performance of their earlier progeny. (The bias is due to a regression to the mean in the subsequent progeny of the selected sires.) However, if the records of progeny from before and after the point at which the selection decision is made are analyzed separately, both may provide an unbiased estimate of the genetic trend (Smith, 1962). An alternative approach, used by Syrstad (1966), is to adjust the estimate of the genetic trend by an estimate of the bias resulting from sire selection. It is worth noting that Smith's method for estimating genetic trend is very similar to that proposed by Van Vleck and Henderson (1961b). Van Vleck and Henderson estimated genetic change using weighted contemporary deviations of A.I. compared with non-A.I. bred animals. If the deviations are calculated on a within-herd x sire subclass basis and regressed on time, the methods are equivalent.

Some of the estimates of genetic trend that have been obtained from the application of the technique of Smith (1962) have already been noted (Acharya and Lush, 1968; Burnside and Legates, 1967; and Legates, 1971). Using Smith's technique, Arave et al. (1964) estimated the genetic trend in fat-corrected milk yield in a single California Jersey herd in the period 1932 to 1962 (1,692 mature-equivalent records) to be 0.7 percent of the herd mean per year. Similarly, Syrstad (1966) obtained an estimate of the genetic trend in milk yield in Norwegian cattle from the records of cows milked in the period 1957 to 1964 (55,200 records of 3-, 4- and 5-year-old cows) of 0.6 percent of the population mean per year. After pooling the estimates of genetic trend produced by three variants of Smith's method, Harville and Henderson (1967) estimated the genetic trend in milk and milkfat production of Holsteins milked in New York State in the period 1956 to 1962 to be 0.75 and 0.65 percent of the mean per year, respectively. Hillers and Freeman (1966) also used Smith's technique and estimated the genetic trend in milk and milkfat production from the records of cows milked in 76 California dairy herds in the period 1951 to 1962 to be 0.26 and 0.24 percent of the mean, respectively. The application by Gustafson (1970) of Smith's method to the records of 119,003 Swedish cattle (of all ages) milked in the period 1959 to 1969 produced an estimate of the genetic trend in fat-corrected milk yield of approximately 1.0 percent of the mean per year. Hargrove and Legates (1971) studied first-lactation records of cows milked in South-eastern regions of the United States in the period 1964 to 1969 and, using Smith's technique, estimated the genetic trend in milk production, as a percentage of breed mean, for Holsteins and Jerseys to be 0.84 and 0.63, respectively. The corresponding estimates for milkfat yield were 0.64 (Holsteins) and 0.45 (Jerseys).

The estimates of genetic trends actually achieved in dairy cattle populations cited in this review are consistently lower than those that have been predicted from a knowledge of the genotypic and phenotypic parameters pertaining to individual selection programmes. This feature of failing to achieve expectation in genetic trend, has been demonstrated in selection studies with other species, for example, mice (Falconer, 1955) although, Chapman (1973) noted that there has been reasonable agreement between estimates of actual and expected genetic trend in single-trait selection studies with sheep. As far as commercial dairy cow populations are concerned, a possible reason for failing to achieve the expected estimates of genetic trend relates to the fact that productive performance may not be the sole criterion that is used by farmers in making selection decisions. The study of P. D. Miller (1969) points to the delay between conception and the time when progeny have their first lactation as being a major part of the reason why expected and actual estimates of genetic trend may not coincide. Predicted estimates of genetic trend are based on estimated genetic selection differentials in the parental generation, whereas estimates of actual genetic trend are based on progeny performance. In a dairy cow population where genetic trend does occur, the period of four years between conception and the completion of the first lactation of the heifer means that, relative to a predicted estimate of the genetic trend, an estimate of genetic trend actually achieved which is based on progeny performance is effectively four years "out-of-date". In a similar vein, Van Vleck and Carter (1972) noted that estimates of daughter superiority (genetic value) obtained from pedigree information were ". . . substantially higher . . . " than estimates based on the daughter's performance.

# CHAPTER 3

# METHODS

## I. ESTIMATION OF GENETIC TREND

Estimates of genetic trend were obtained from the application of two different methods both of which are variants of the technique originally proposed by Smith (1962). The mathematical description of Smith's approach is:

$$\text{Genetic trend} = -2\, b_{(P-S)T} \qquad \ldots (1)$$

where: $b$ = the pooled within-sire regression coefficient.

$(P-S)$ = the deviation of the progeny average (S) from the population average (P), obtained on a within-year basis to overcome the effects of year-to-year variation in the environment on the estimate of the genetic trend.

$T$ = time (year) in which the deviation was calculated.

In populations where the genetic trend is positive, the sign of the regression coefficient in equation (1) is negative; hence the need for the minus sign in the right-hand-side of equation (1).

In data derived from a multi-herd dairy cow population, environmental variation from herd-to-herd as well as year-to-year, if not taken into account will effect the accuracy of the estimate of genetic trend. If the population average in equation (1) is replaced by a herdmate average, so that the deviations are then calculated on a within-year, within-sire x herd subclass basis, the effects of both these sources of environmental variation on the estimate of genetic trend are minimized. However, if the herdmate averages are derived from the records of cows of all age groups, the data must be age corrected. In Chapter 2, it was noted that inaccurate age corrections or the culling of females are both sources of bias when second and later lactation records are included in herdmate averages. Since the present study involved the analysis of first-lactation records only, these sources of bias were automatically eliminated.

## A. Method One

$$\text{Genetic trend/year} = -2b_{(C-S)T}$$

$$= -2\left[\frac{\sum_i \sum_j \left[\sum_k d_{ijk}t_{ijk} - \dfrac{\left(\sum_k d_{ijk}\right)\left(\sum_k t_{ijk}\right)}{n}\right]}{\sum_i \sum_j \left[\sum_k t_{ijk}^2 - \dfrac{\left(\sum_k t_{ijk}\right)^2}{n}\right]}\right] \quad . . . (2)$$

where: $d_{ijk}$ = the average deviation of first-lactation daughters in the ijth sire x herd subclass from the average of their contemporaries (C) in the kth year.

$t_{ijk}$ = the kth year in which the deviation of daughters in the ijth sire x herd subclass was calculated.

$n$ = the number of years in which deviations of the ijth subclass were able to be calculated.

The standard error of an estimate of genetic trend produced by the application of equation (2) was calculated using the formula:

$$\text{Std. Error} = 2\sqrt{\frac{1}{N-2}\left[\frac{\sum_i\sum_j\left[\sum_k d_{ijk}^2 - \left(\frac{\sum_k d_{ijk}}{n}\right)^2\right]}{\sum_i\sum_j\left[\sum_k t_{ijk}^2 - \left(\frac{\sum_k t_{ijk}}{n}\right)^2\right]} - b^2_{(C-S)T}\right]}$$

where: $b_{(C-S)T}$, $d_{ijk}$, $t_{ijk}$ and n are as defined in equation (2)

and where: N = the sum of n over all sire x herd subclasses.

Hickman (1971) noted that the application of method one to estimate genetic trend " . . . is complicated by the fact that the actual and relative number of daughters and contemporaries varies over time." Consequently, he suggested the employment of the weighting procedure devised by Robertson et al. (1956) to prevent daughter deviations from large or small numbers of contemporaries having an undue effect on the regression coefficient. The appropriate within-year and sire x herd subclass weighting factor derived by Robertson et al.$^{(a)}$ is:

$$w_{ijk} = \left(\frac{n_{C_{ijk}} \times n_{S_{ijk}}}{n_{C_{ijk}} + n_{S_{ijk}}}\right)$$

---

$^{(a)}$ Variance ($d_{ijk}$) = Variance $\left(\frac{S_{ijk}}{n_{S_{ijk}}} - \frac{C_{ijk}}{n_{C_{ijk}}}\right)$

which, if the two groups are independent: $= \left(\frac{n_{S_{ijk}} + n_{C_{ijk}}}{n_{S_{ijk}} \times n_{C_{ijk}}}\right)$ (Variance of individual records.)

where: $S_{ijk}$ and $C_{ijk}$ are the sums of daughter and contemporary yields in the kth year respectively. Since the variance of individual records is constant, the inverse of the variance ($d_{ijk}$) becomes $w_{ijk}$ as given in the text.

where: $w_{ijk}$ = the inverse of the variance of the deviation $d_{ijk}$

$n_{C_{ijk}}$ = the number of contemporaries in the jth herd to which the daughters of the ith sire were compared in the kth year (excluding the records of the sire's daughters)

$n_{S_{ijk}}$ = the number of daughters of the ith sire in the jth herd and kth year on which the average deviation $d_{ijk}$ is based.

Therefore, Hickman proposed that the genetic trend be estimated using a pooled within-sire regression of weighted daughter deviations from contemporary averages on time (as opposed to the unweighted deviations used in method one).

## B. Method Two

Genetic trend/year = $-2wb_{(C-S)T}$

$$= -2 \left[ \frac{\sum_i \sum_j \left[ \sum_k w_{ijk} d_{ijk} t_{ijk} - \frac{\left( \sum_k w_{ijk} d_{ijk} \right) \left( \sum_k \right.}{\sum_k w_{ijk}} \right]}{\sum_i \sum_j \left[ \sum_k w_{ijk} t_{ijk} - \frac{\left( \sum_k w_{ijk} t_{ijk} \right)^2}{\sum_k w_{ijk}} \right]} \right] \quad 3$$

where: $w_{ijk}$, $d_{ijk}$ and $t_{ijk}$ are as previously defined in the text.

The standard error of an estimate of genetic trend produced by the application of equation (3) was calculated using the formula:

$$\text{Std. Error} = 2\sqrt{\frac{1}{\sum_{ijk} w_{ijk} - 2}\left[\frac{\sum_i\sum_j\left[\sum_k w_{ijk}d_{ijk}^2 - \frac{\left(\sum_k w_{ijk}d_{ijk}\right)^2}{\sum_k w_{ijk}}\right]}{\sum_i\sum_j\left[\sum_k w_{ijk}t_{ijk}^2 - \frac{\left(\sum_k w_{ijk}t_{ijk}\right)^2}{\sum_k w_{ijk}}\right]} \quad (C-S)T\right]}$$

where: $wb_{(C-S)T}$, $w_{ijk}$, $d_{ijk}$ and $t_{ijk}$ are as previously defined.

If the sires have been selected for breeding on the basis of the productive performance of their earlier progeny, the estimates of genetic trend obtained from the application of either method one or method two, will be biased. In line with the suggestion of [illegible] (1962), part of the bias was eliminated by excluding from the analysis the records of progeny sired prior to the point at which [illegible] decision is usually made. Therefore, the records of [illegible] in sire-proving herds were disregarded since it is these records which largely determined whether or not the bulls represented in this study were included in the proven bull team for use in A.I. In the case of non-A.I. sires, the bias due to sire selection is not likely to be important. Even though a bull has entered the proven bull [illegible] sire selection still occurs through a progressive type of [illegible] on the basis of progeny performance to allow new bulls to enter the team. The bias due to this cause is difficult to eliminate.

Estimates of genetic trend produced by method one and method two are biased by non-random mating of sires and dams, particularly across cow age groups (e.g., older sires mated to older cows). The bias has already been discussed in Chapter 2. If the genetic trend estimates are based on the regression of a within-cow age and within-sire x herd subclass deviation on time, the bias due to non-random mating in dairy herds will, to a large extent, be overcome although this was not attempted in this study.

The pooled intra-sire regression can only be calculated if individual sires are represented by daughter deviations from contemporary averages in at least three years. If the deviations for a particular sire are obtained from a different herd each year, the estimate of the genetic trend is confounded with genetic differences between the herds. This problem can be overcome, as in this study, by requiring a sire to be represented within individual herds on at least three occasions so that the genetic trend is estimated on the basis of a pooled within-sire, within-herd regression of daughter deviations on time.

## II. SIRE EVALUATION BASED ON A MIXED MODEL

In this section a technique which provides an estimate of the genetic merit of individual sires, together with their subsequent ranking, and which is unbiased by genetic trend or differential average merit of the herdmates to which the daughters of [illegible] are compared, is described. A general explanation of the method was given by Lentz et al. (1969) whilst its details have been elucidated by Henderson (1973) and Keown (1974). In this study, the model upon which the method is based was made more appropriate to the New Zealand situation and was applied to the edited-A data (defined in Chapter 4).

The model used was:

$$Y_{ijkl} = g_i + s_{ij} + h_k + e_{ijkl}$$

where: $Y_{ijkl}$ = the first-lactation record made in the kth herd year by the lth daughter of the jth sire belonging to the ith sire-group.

$g_i$ = the effect of the ith sire-group (taken to include the overall mean $\mu$) where sire grouping is based on the year at which it was first represented by a daughter with a lactation record. Group effects assumed to be fixed effects.

$s_{ij}$ = the effect of the jth sire of the ith sire-group. Sire effects are assumed to be random effects distributed with zero mean and variance $I\sigma_s^2$

$h_k$ = the effect of the kth herd-year. Herd-year effects are considered to be random effects distributed with mean zero and variance $I\sigma_h^2$

$e_{ijkl}$ = the random error effect unique to the lth daughter of the jth sire in the ith sire-group in herd-year k. Error effects are distributed with mean zero and variance $I\sigma_e^2$

It is assumed that the variables in the model are uncorrelated.

Through the grouping of sires, the breeding value of each sire is estimated relative to those which belong to the same group. Therefore, by arranging for sires to be grouped according to some time scale (in this case, time of first usage), the bias due to genetic trend is eliminated. The strength of the mixed model approach for sire evaluation is that sires are evaluated on the basis of direct and indirect comparisons with one another on an individual basis and the use of herdmate averages or variations thereof is avoided. Therefore, biases due to factors such as unequal gene [illegible] the herdmates of daughters of the sires are automatica[illegible]

The mixed model A can be re-written in matrix notation:

$$Y = Xb + Zu + e$$

where:

Y is a known N x 1 vector of first-lactation [illegible] records. (N is the total number [illegible])

X is a known full-rank N x p incidence matrix of the fixed effects of the model.

| | |
|---|---|
| b | is an unknown p x 1 vector of the group effects. |
| Z | is a known non-full-rank N x q incidence matrix of the random effects of the model. (q is the combined number of sire and herd-year effects.) |
| u | is an unknown q x 1 vector of the sire and herd-year effects of the model. |
| e | is an unknown N x 1 vector of error effects also assumed to be random effects. |

The appropriate expected values for this model are:

$E(u) = E(e) = 0$ (u and e are vectors of random effects.)

$\therefore E(Y) = Xb$

$E(ue') = 0$ (u and e are assumed to be uncorrelated.)

$E(be') = 0$ (b and e' are assumed to be uncorrelated.)

$E(uu') = D\sigma_e^2$

where: $D\sigma_e^2$ is the variance-covariance matrix of the random effects of the model apart from error effects.

$E(ee') = R\sigma_e^2$

where: $R\sigma_e^2$ is the variance-covariance matrix of the error effects of the model.

It is not difficult to demonstrate[(b)] that the variance-covariance matrix of the observations (lactation records) represented by the vector Y is: $(ZDZ' + R)\,\sigma_e^2$

If the Method of Mixed Model Solution (MMS) (Henderson, 1973) is applied to model (B), the resultant normal equations are:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & (Z'R^{-1}Z + D^{-1}) \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}Y \\ Z'R^{-1}Y \end{bmatrix} \qquad \ldots (4)$$

where: $\hat{b}$ and $\hat{u}$ are estimates of b and u, respectively.

All other matrices are as previously defined.

The salient details on which the derivation of equations (4) are based, are given in the next sub-section.

---

(b) Variance-covariance (Y) $= E\left[Y-E(Y)\right]\left[Y-E(Y)\right]'$

$$= E\left[Zu + e\right]\left[Zu + e\right]'$$

$$= E\left[Zuu'Z' + eu'Z' + Zue' + ee'\right]$$

$$= E\left[Zuu'Z'\right] + E\left[ee'\right]$$

$$= (ZDZ' + R)\,\sigma_e^2$$

## A. The Method of Mixed Model Solution

Given the vector Y, the desired prediction is a set of s linear functions:

$$K'b + M'u + V'f \qquad \ldots (5)$$

where: b and u are as defined.

K' is a known s x p matrix.

M' is a known s x q matrix.

V' is a known s x N matrix.

f is a non-observable random vector with E(f) = 0 and variance-covariance matrix $R_1\sigma^2_e$. Furthermore, f is uncorrelated with u and e.

Equation (5) is referred to as the predictand.

The prediction is limited to predictors which are linear functions of Y, for example, T'Y where T' is an s x N matrix to be determined. Following Henderson (1973), the criterion on which the determination of the matrix T is based is that the average squared prediction error be minimized.

Now: Prediction error = Predictor - Predictand

$$= (T'Y - K'b - M'u - V'f).$$

Then the average squared prediction error is:

$$E(T'Y - K'b - M'u - V'f)\,(T'Y - K'b - M'u - V'f)' \qquad \ldots (6)$$

The expansion of (6), coupled with the removal of the expressions with zero expectations, yields:

$$(T'X - K')\,bb'(T'X - K')' + (T'Z - M')D(T'Z - M')'\sigma^2_u +$$
$$(T'RT)\sigma^2_e + (V'R_1V)\,\sigma^2_e \qquad \ldots (7)$$

Therefore, the required value of T is that which minimizes the diagonal elements of (7) and this can be achieved by differentiating that equation with respect to T followed by setting the resultant expression to zero.

That is: $\left[ Xbb'X' + (ZDZ' + R)\sigma_e^2 \right] T - Xbb'K - ZDM\sigma_e^2 = 0$

so that: $\left[ Xbb'X' + (ZDZ' + R)\sigma_e^2 \right] T = Xbb'K + ZDM\sigma_e^2 \quad \ldots (8)$

Unless the parameters b and $\sigma_e^2$ are known (which is not usually the case), equation (8) cannot be solved. However, a convenient way of overcoming this predicament is to require the prediction to be unbiased.

That is: E(Predictor) = E(Predictand).

But: E(Predictand) = K'b,

while: E(Predictor) = T'E(Y) = T'Xb.

Consequently, the prediction is unbiased when $T'X = K'$ (or, $K = X'T$).

The requirement that the prediction be unbiased means that the average squared prediction error, as represented in equation (7), can be reduced to:

$$(T'Z - M')D(T'Z - M')'\sigma_e^2 + T'RT\sigma_e^2 + V'R_1V\sigma_e^2 \quad \ldots (9)$$

because: $(T'X - K')bb'(T'X - K')' = 0$

Now, differentiating (9) with respect to T and with the restriction to $X'T = K$, yields the equations:

$$\begin{bmatrix} (ZDZ' + R) & X \\ X' & 0 \end{bmatrix} \begin{bmatrix} T \\ \theta \end{bmatrix} = \begin{bmatrix} ZDM \\ K \end{bmatrix}$$

where: $\theta$ is a set of Lagrange multipliers.

$\sigma_e^2$ is a constant and has been eliminated.

Although (10) can be solved, in most situations the computational requirements are prohibitive. An alternative approach is to regard u as a vector of fixed effects. Under these conditions, the term (ZDZ' + R) reduces to R since D (and therefore, ZDZ') is zero.

Starting with equation (6), the steps outlined so far could be reworked, but this is unnecessary since (10) can be used as a basis for an analagous set of equations. If u is taken to be fixed, two restrictions on the minimization of the average squared prediction error equation are required.

That is: $X'T = K$ and, $Z'T = M$.

Equations (10) can now be modified to:

$$\begin{bmatrix} R & X & Z \\ X' & 0 & 0 \\ Z' & 0 & 0 \end{bmatrix} \begin{bmatrix} T \\ \Theta_1 \\ \Theta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ K \\ M \end{bmatrix} \qquad . . . (11)$$

where: $\Theta_1$ is a set of Lagrange multipliers pertaining to the restriction $X'T = K$.

$\Theta_2$ is a set of Lagrange multipliers pertaining to the restriction $Z'T = M$.

However, model (B) requires that u be regarded as a vector of random effects. If $D^{-1}$ is added to (11) as indicated below, the fact that u is random is taken into account.

That is:

$$\begin{bmatrix} R & X & Z \\ X' & 0 & 0 \\ Z' & 0 & -D^{-1} \end{bmatrix} \begin{bmatrix} T \\ \Theta_1 \\ \Theta_2 \end{bmatrix} = \begin{bmatrix} 0 \\ K \\ M \end{bmatrix} \qquad . . . (12)$$

It is not difficult to demonstrate[(c)] that equations (12) are equal to equations (10) (which were obtained through taking u to be random from the beginning of the exercise). In contrast to the situation with regard to (10), the solution of (12) is readily obtainable.

Consider now the first equation in (12). Re-organization of that equation produces:

$$T = R^{-1}X\theta_1 - R^{-1}Z\theta_2 \qquad \ldots (13)$$

Substitution of this result into the remaining equations of (12) yields:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & (Z'R^{-1}Z+D^{-1}) \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = - \begin{bmatrix} K \\ M \end{bmatrix} \qquad \ldots (14)$$

If $\begin{bmatrix} P & Q \\ Q' & S \end{bmatrix}$ is any symmetric generalized inverse of the coefficient matrix of equations (14) then; $\theta_1 = -(PK + QM)$, and $\theta_2 = -(Q'K + SM)$.

Equation (13) can be re-written as:

$$T = R^{-1}X(PK + QM) + R^{-1}Z(Q'K + SM).$$

Therefore, $T'Y = K'(PX'R^{-1}Y + QZ'R^{-1}Y) + M'(Q'X'R^{-1}Y$

---

(c) From the last equation in (12), $\theta_2 = DZ'T - DM$.

Substituting for $\theta_2$ in the first equation in (12) produces:

$$RT + X\theta_1 + ZDZ'T - ZDM = 0.$$

This can then be re-organized to:

$$\begin{bmatrix} (ZDZ' + R) & X \\ X' & 0 \end{bmatrix} \begin{bmatrix} T \\ \theta_1 \end{bmatrix} = \begin{bmatrix} ZDM \\ K \end{bmatrix}$$

An important result is that the generalized inverse of the coefficient matrix of equations (14) is the same as that for equations (4).

Solving (4) for $\hat{b}$ and $\hat{u}$ produces:

$$\hat{b} = (PX'R^{-1}Y + QZ'R^{-1}Y)$$

$$\hat{u} = (QX'R^{-1}Y + SZ'R^{-1}Y).$$

Hence, equations (15) can be written as: $T'Y = K'\hat{b} + M'\hat{u}$, which, apart from V'f, represents the original prediction equation with $\hat{b}$ and $\hat{u}$ substituted for b and u, respectively. Therefore, best linear unbiased estimation of fixed effects and best linear unbiased prediction of random effects in a mixed model is achieved by solving equations (4). The main point about these equations, however, is that they are simply modified least squares equations. In a model where all effects are assumed fixed but where the effects have been partitioned into the two vectors b and u, the normal equations produced by the application of the method of least squares are:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}Y \\ Z'R^{-1}Y \end{bmatrix}$$

The fact that in a mixed model u is a vector of random as opposed to fixed effects, can be taken account of through simply modifying ordinary least squares equations by adding $D^{-1}$ as indicated in equations (4).

## B. Adaptation of the Mixed Model Solution Equations to the Sire Evaluation Model

The definition of the sire evaluation model, that is, model [illegible], included a number of assumptions. These were:

(a) Sire effects are uncorrelated so that:

Covariance $(s_{ij}, s_{ij'}) = 0$ and therefore,
Variance $(s_{ij}) = I\sigma_s^2$ where $\sigma_s^2$ is the sire variance and I is the identity matrix.

(b) Herd-year effects are uncorrelated so that:

Covariance $(h_k, h_{k'}) = 0$ and therefore,
Variance $(h_k) = I\sigma_h^2$ where $\sigma_h^2$ is the herd-year variance.

(c) Sire and herd-year effects are uncorrelated so that:

Covariance $(s_{ij}, h_k) = 0$.

(d) Error effects are uncorrelated and consequently:

Covariance $(e_{ijkl}, e_{ijkl'}) = 0$ so that:
Variance $(e_{ijkl}) = I\sigma_e^2$ where $\sigma_e^2$ represents the error variance.

The assumptions (a), (b) and (c) have important implications as far as the variance-covariance matrix $D\sigma_e^2$ is concerned. Recalling that $D\sigma_e^2 = E(uu')$, these assumptions allow $D\sigma_e^2$ to be written in the form:

$$E(uu') = \begin{bmatrix} \sigma_u^2 & & & 0 \\ & \cdot & & \\ 0 & & \cdot & \\ & & & \sigma_u^2 \end{bmatrix} = I\sigma_u^2 \text{ where}$$

$\sigma_u^2$ represents the variance of the random effects in the model.

Since $D\sigma_e^2 = I\sigma_u^2$, D must now $= I\frac{\sigma_u^2}{\sigma_e^2}$

For the diagonal elements of D which correspond to sire effects, $\sigma^2_u = \sigma^2_s$ and likewise, for the diagonal elements of D which correspond to herd-year effects, $\sigma^2_u = \sigma^2_h$. Therefore, the matrix D is simply a diagonal matrix of variance ratios. In turn, the inverse matrix $D^{-1}$ is easily obtained by taking reciprocals of the diagonal elements.

Assumption (d) allows the variance-covariance matrix of the error effects ($R\sigma^2_e$) to be reduced to $I\sigma^2_e$. That being the case, the appropriate form of equations (1) for the sire evaluation model that has been described are now:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & (Z'Z+D^{-1}) \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'Y \\ Z'Y \end{bmatrix} \quad \ldots (16)$$

In this study, the breeding values of the sires were estimated from a solution of equations (16).

## C. Variance Component Estimation

The solution of the mixed model equations (equations 16, depends upon the prior estimation of the variance components of the model which, in this case, are the sires, herd-year and error variances. When data are balanced, estimates of variance components are readily obtainable by equating mean squares from an analysis of variance to their expected values. However, in the case of unbalanced data (as in this study) and, more especially, where the model is a mixed model, the estimation of variance components is not simple.

A major contribution to the problem of estimating variance components for mixed models with unbalanced data was made by Henderson (1953). He proposed three methods for estimating variance components. In a review and re-statement of Henderson's methods using matrix terminology, Searle (1968) showed that Henderson's Method II was not uniquely specified and that this deficiency detracted from its value as a useful procedure.

Henderson's Method I, which was used in this study, is simply an analogue of the analysis of variance method used for estimating variance components with balanced data. All effects of the model, including fixed effects in mixed models, are regarded as random effects. In contrast to the situation with balanced data, there is no unique partitioning of sums of squares with unbalanced data and, as far as mixed models are concerned, it is impossible to obtain expectations for random effects which do not contain quadratic functions of the fixed effects of the model. Thus, Henderson's Method I is incapable of yielding variance components which are unbiased by the fixed effects in mixed models. For random models, the method does produce unbiased estimates of variance components.

Unlike Method I, Henderson's Method III does not utilize analysis of variance sums of squares but rather, reductions in sums of squares due to fitting different sub-models. In applying Method III to mixed models, it is through the judicious choice and fitting of sub-models that reductions in sums of squares which do not contain functions of the fixed effects of the model are produced. Unbiased estimates of the variance components can then be obtained by equating each reduction to its expected value under the full model. The major limitation of Henderson's Method III is the large amount of computation required. Obtaining the reductions in sums of squares can involve the inversion of matrices of order equal to the number of random effects in the data which, as far as this study is concerned, would have exceeded the capacity of the computer. The iterative procedure suggested by Cunningham and Henderson (1968), but later corrected by Thompson (1969), was aimed at circumventing the computational requirement of Henderson's Method III in obtaining estimates of the variance components as a preliminary step in the estimation of the fixed effects in a mixed model. The iterative procedure could not be used in this study since it is applicable to only those mixed models having just one random effect.

Although Henderson's Method III is the method of choice in a mixed model, it was not used because of its computational difficulty.

(1) Application of Henderson's Method I. Estimates of the variance components were obtained assuming the following model:

$$Y_{ijkl} = \mu + g_i + s_{ij} + h_k + (sh)_{ijk} + e_{ijkl} \qquad \ldots \text{(C)}$$

where: $Y_{ijkl}$ = the first-lactation record made in the kth herd-year by the lth daughter of the jth sire belonging to the ith sire group.

$\mu$ = a general mean.

$g_i$ = the random effect of the ith sire-group. Group effects are distributed with mean zero and variance $I\sigma_g^2$.

$s_{ij}$ = the random effect of the jth sire in the ith sire-group. Sire effects are distributed with mean zero and variance $I\sigma_s^2$.

$h_k$ = the random effect of the kth herd-year. Herd-year effects are distributed with mean zero and variance $I\sigma_h^2$.

$(sh)_{ijk}$ = the random interaction effect of the ijth sire by kth herd-year subclass. Sire x herd-year effects are distributed with mean zero and variance $I\sigma_{hs}^2$.

$e_{ijkl}$ = the random error effect unique to the lth daughter of the jth sire in the ith sire-group in herd-year k. Error effects are distributed with mean zero and variance $I\sigma_e^2$.

It is assumed that the variables of the model are independent of each other and that the other possible interaction term, the group by herd-year interaction, is zero.

The initial step in applying Henderson's Method I is to obtain expectations of the uncorrected analysis of variance sums of squares. These are (using S.S. to denote uncorrected sums of squares):

$$E\left[S.S.\right] = E\begin{bmatrix}\text{Total S.S.}\\ \text{Group S.S.}\\ \text{Sire within group S.S.}\\ \text{Herd-year S.S.}\\ \text{Herd-year x Sire within group S.S.}\\ \text{Herd-year x Group S.S.}\\ \text{Correction term}\end{bmatrix}$$

$$= E\begin{bmatrix}\sum_i\sum_j\sum_k\sum_l Y^2_{ijkl}\\ \sum_i Y^2_{i...}/n_{i..}\\ \sum_i\sum_j Y^2_{ij..}/n_{ij.}\\ \sum_k Y^2_{..k.}/n_{..k}\\ \sum_i\sum_j\sum_k Y^2_{ijk.}/n_{ijk}\\ \sum_i\sum_k Y^2_{i.k.}/n_{i.k}\\ Y^2_{....}/N\end{bmatrix}$$

where: The herd-year x group interaction sum of squares has been included so that the expectation of the corrected herd-year x sire within group interaction sums of squares can be computed.

$n_{ijk}$ is the number of daughters of the jth sire in ith sire-group in herd-year k.

N is the total number of records.

The expanded form of $E\left[\text{S.S.}\right]$ is:

$$\begin{bmatrix}
N & N & N & N & N & 0 & N \\
N & N & \sum_i\left[\frac{\sum_j n_{ij.}^2}{n_{i..}}\right] & \sum_i\left[\frac{\sum_k n_{i.k}^2}{n_{i..}}\right] & \sum_i\left[\frac{\sum_j\sum_k n_{ijk}^2}{n_{i..}}\right] & 0 & a \\
N & N & N & \sum_i\sum_j\left[\frac{\sum_k n_{ijk}^2}{n_{ij.}}\right] & \sum_i\sum_j\left[\frac{\sum_k n_{ijk}^2}{n_{ij.}}\right] & 0 & b \\
N & \sum_k\left[\frac{\sum_i n_{i.k}^2}{n_{..k}}\right] & \sum_k\left[\frac{\sum_i\sum_j n_{ijk}^2}{n_{..k}}\right] & N & \sum_k\left[\frac{\sum_i\sum_j n_{ijk}^2}{n_{..k}}\right] & 0 & c \\
N & N & N & N & N & 0 & f \\
N & N & \sum_i\sum_k\left[\frac{\sum_j n_{ijk}^2}{n_{i.k}}\right] & N & \sum_i\sum_k\left[\frac{\sum_j n_{ijk}^2}{n_{i.k}}\right] & 0 & m \\
N & \left[\frac{\sum_i n_{i..}^2}{N}\right] & \left[\frac{\sum_i\sum_j n_{ij.}^2}{N}\right] & \left[\frac{\sum_k n_{..k}^2}{N}\right] & \left[\frac{\sum_i\sum_j\sum_k n_{ijk}^2}{N}\right] & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\mu^2 \\ \sigma_g^2 \\ \sigma_s^2 \\ \sigma_h^2 \\ \sigma_{hs}^2 \\ \sigma_{hg}^2 \\ \sigma_e^2
\end{bmatrix} = CV$$

where: $C$ = the coefficient matrix.

$V$ = the column vector of variance components.

$a$ = the number of sire groups.

$b$ = the total number of sires.

$c$ = the number of herd-years.

$f$ = the number of herd-year x sire within group subclasses containing non-zero data.

$m$ = the number of herd-year x group subclasses containing non-zero data.

and where: the column of zeros in the matrix C represents the assumption of the model that the herd-year x group interaction variance ($\sigma^2_{hg}$) is zero.

Having obtained the expectations of the uncorrected sums of squares, it is necessary to obtain the expectations of the corrected sums of squares. That is:

$$\begin{bmatrix} \text{E(Group S.S.) - E(Correction term)} \\ \text{E(Sire within group S.S. - E(Group S.S.)} \\ \text{E(Herd-year S.S.) - E(Correction term)} \\ \begin{array}{l} \text{E(Herd-year x Sire within group S.S.)} \\ \text{- E(Sire within group S.S.)} \\ \text{- E(Herd-year x group S.S.) + (Group S.S.)} \end{array} \\ \text{E(Total S.S.) - E(Herd-year x Sire within group S.S.)} \end{bmatrix} = A \begin{bmatrix} \sigma^2_g \\ \sigma^2_s \\ \sigma^2_h \\ \sigma^2_{hs} \\ \sigma^2_e \end{bmatrix}$$

where: $A$ = the coefficient matrix obtained by subtracting and adding rows of the coefficient matrix C in accordance with the way in which corrected sums of squares are computed.

If the column vector of the analysis of variance corrected sums of squares is W, estimates of the variance components can be obtained as follows:

$$A^{-1}W = \begin{bmatrix} \hat{\sigma}^2_g \\ \hat{\sigma}^2_s \\ \hat{\sigma}^2_h \\ \hat{\sigma}^2_{hs} \\ \hat{\sigma}^2_e \end{bmatrix}$$

## D. Computation of the Solution to the Mixed Model Equations

The solution of the mixed model equations (equations 16) involves the calculation of the inverse of the matrix $(Z'Z + D^{-1})$. The order of this matrix is equal to the number of random effects in the data and therefore, obtaining its inverse may be difficult from a computational standpoint. A convenient way around this problem is to detail the structure of the mixed model equations in a way which relates to further partitioning.

Recall equations (16):

$$\begin{bmatrix} X'X_{(p \times p)} & X'Z_{(p \times q)} \\ Z'X_{(q \times p)} & (Z'Z + D^{-1})_{(q \times q)} \end{bmatrix} \begin{bmatrix} \hat{b}_{(p \times 1)} \\ \hat{u}_{(q \times 1)} \end{bmatrix} = \begin{bmatrix} X'Y_{(p \times 1)} \\ Z'Y_{(q \times 1)} \end{bmatrix}$$

where: $p$ = number of fixed effects (groups).

$q$ = number of random effects (sires + herd-years).

The matrices which are functions of the matrix Z can be partitioned into sub-matrices corresponding to the sire and the herd-year effects.

That is:

$$\begin{bmatrix} X'X_{(p \times p)} & E_{(p \times e)} & F_{(p \times c)} \\ E'_{(e \times p)} & S_{(e \times e)} & B_{(e \times c)} \\ F'_{(c \times p)} & B'_{(c \times e)} & H_{(c \times c)} \end{bmatrix} \begin{bmatrix} \hat{b}_{(p \times 1)} \\ \hat{s}_{(e \times 1)} \\ \hat{h}_{(c \times 1)} \end{bmatrix} = \begin{bmatrix} X'Y_{(p \times 1)} \\ G_{(e \times 1)} \\ Q_{(c \times 1)} \end{bmatrix}$$

where: $e$ = the number of sires.

$c$ = the number of herd-years.

$S$ = a diagonal matrix of the number of daughters of each sire with the variance ratio $\frac{\hat{\sigma}^2_e}{\hat{\sigma}^2_s}$ added.

$H$ = a diagonal matrix of the number of records of each herd-year with variance ratio $\frac{\hat{\sigma}^2_e}{\hat{\sigma}^2_h}$ added.

$\hat{s}$ = a vector of sire effects.

$\hat{h}$ = a vector of herd-year effects.

$G$ = a vector of sire totals ($Y_{.j..}$).

$Q$ = a vector of herd-year totals ($Y_{..k.}$).

and where: E and F are sub-matrices of X'Z.

B is a matrix of the number of daughters of each sire in each herd-year.

Since it is unnecessary for herd-year effects to be estimated, the equations for herd-year effects are absorbed. The equations to be solved then become:

$$\begin{bmatrix} (X'X - FH^{-1}F') & (E - FH^{-1}B') \\ (E - FH^{-1}B')' & (S - BH^{-1}B') \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{s} \end{bmatrix} = \begin{bmatrix} (X'Y - FH^{-1}Q) \\ (G - BH^{-1}Q) \end{bmatrix} \quad \ldots (17)$$

with solutions:

$\hat{b}$ = solution for the group effects

$$= \left[(X'X - FH^{-1}F') - (E - FH^{-1}B')(S - BH^{-1}B')^{-1}(E - FH^{-1}B')'\right]^{-1} \times \left[(X'Y - FH^{-1}Q) - (E - FH^{-1}B')(S - BH^{-1}B')^{-1}(G - BH^{-1}Q)\right]$$

$\hat{s}$ = estimate of the sire effects

$$(S - BH^{-1}B')^{-1} \left[(G - BH^{-1}Q) - (E' - BH^{-1}F')\hat{b}\right]$$

Therefore, the order of the largest non-diagonal matrix to be inverted has been reduced by the number of herd-years down to the number of sires in the study.

According to the sire evaluation model, the effect of a sire is estimated relative to the effects of only the sires in the same sire group. Therefore, sires cannot be compared across groups until the group effects have been taken into account, i.e., until $\hat{g}_i + \hat{s}_{ij}$ has been calculated for all sires. The quantity $\hat{g}_i + \hat{s}_{ij}$ represents the estimated breeding value of the jth sire in the ith sire-group.

## E. Re-definition of the Sire Evaluation Model

In the original sire evaluation model (model A), the herd-year effects were assumed to be random effects. Since individual sires are not represented in all herd-years, it is possible for variation in the average genetic merit of the herds with which each sire is associated to bias the estimates of sire breeding values that are obtained from the application of the mixed model (Henderson, 1973). If, however, herd-year effects are regarded as being fixed effects, the bias is overcome (Henderson, 1973).

Regarding herd-year effects as fixed involves only minor modifications to equations (17). These are:

(i) The application of one further constraint on the solution (in this case, one of the herd-year effects was set to zero).

(ii) The variance component ratio $\frac{\hat{\sigma}^2_e}{\hat{\sigma}^2_h}$ is not added to the diagonal elements of the matrix H.

As before, the herd-year effects are absorbed in the process of computing a solution to the mixed model equations.

In this study, the extent of the bias due to non-random usage of sires amongst herd-years was assessed through a comparison of the estimates of sire breeding values assuming the herd-year effects to be fixed effects on the one hand and random effects on the other.

## F. Genetic Trend Estimation Based on the Mixed Model Solution

Since the sires were grouped according to their year of first usage, $\hat{g}_i$ is the unweighted average genetic merit of the sires sampled in the ith year. Therefore, the trend in $\hat{g}_i$ over time is the genetic trend in the population (Henderson, 1973). The weighted average genetic merit of the sires belonging to group i is given by:

$$\frac{\sum_{j} n_{ij.} (\hat{g}_i + \hat{s}_{ij})}{n_{i..}}$$

Regressing the weighted group effects on time provides an estimate of genetic trend attributable to the progeny of the sires.

In this study, both the weighted and the unweighted group effects were regressed on time to obtain estimates of the genetic trend. These methods for estimating genetic trend were applied only to the edited-A.I. data.

# CHAPTER 4

# DATA

## I. SOURCE OF DATA

A requirement of the methods of analysis used is that the sires of the cows which produced the lactation records be identified. Since the accuracy of sire identification is likely to be highest in those herds where all cows are sire-identified, it was decided that only such herds would be considered as potential contributors of data for this study. In addition, the procedure for estimation of genetic trend required the herds to have been herd-tested for a minimum of three years and that in each herd there must be at least one sire with recorded daughters in all three years. Due to a limitation on resources, it was further decided that the study be restricted to data derived from:

(i) Herds within the Auckland Herd Improvement Association (now Auckland Livestock Improvement Association) and;

(ii) Animals of the Jersey breed (but not Jersey cross).

Each year, the Farm Production Division of the N.Z. Dairy Board lists those herds where it has been possible to compute a within-herd production comparison of sire-identified A.I. with non-A.I. bred animals. It was from the list compiled at the end of the 1971-72 dairying season that the 89 herds which contributed data to this study were identified. Having identified the herds which met those requirements, lactation records were collected from the annual herd-testing summaries of the dairying seasons 1966-67 to 1972-73, inclusive. The numbers of suitable lactation records in the seasons prior to 1966-67 were too small to justify their inclusion.

## II. DESCRIPTION OF DATA

The screening technique employed for sampling these data closely followed that currently used by the Farm Production Division of the N. Z. Dairy Board in collecting data for its sire-evaluation programme. Accordingly, only first-lactation records were sampled. The suitability of each lactation record was determined on the basis of the criteria listed in Table 4.1.

The details collected on individual lactations deemed to have met the requirements of the study were:

(i) Sire code.

(ii) Herd code.

(iii) Year (dairying season) in which the lactation was made.

(iv) The total milk produced in the lactation.

(v) The total milkfat produced in the lactation.

Since the data collected were all first-lactation records, there was no necessity for the records to be corrected for age. In line with the policy of the Farm Production Division of the N. Z. Dairy Board, there was no general adjustment of the data for variation in lactation length. It has already been noted that lactations shorter than 100 days in length were excluded. At the other end of the scale,

Table 4.1 : Criteria determining the suitability of the lactation records.

---

**Exclusions:**

- (a) Records from cows milked for less than 100 days.
- (b) Records from cows milked for between 100 and 200 days but culled because of:
    - (1) Injury (including lameness).
    - (2) Disease (abortion, bloat, Brucellosis, Johne's disease, Leptospirosis, mastitis, metabolic diseases, milk fever, Salmonella or Tuberculosis).
- (c) Records from cows milked for between 100 and 200 days but died of catarrh.
- (d) Records from cows milked in sire-proving herds.

**Inclusions:**

Records from any cows milked for 100 days or more including those subsequently culled because of:

- (1) Low production.
- (2) Low fertility (including calving difficulty)
- (3) Other diseases.

---

records from lactations exceeding 305 days in length were corrected back to the equivalent of that length by deduction of a proportion of the cow's production during the lactation. A description of the data collected for this study is given in Table 4.2.

Each lactation record was sire-identified by one of three classes of sire code. These classes were:

(i) An A.I. sire code.

(ii) A Non-A.I. (registered) sire code.

(iii) No sire code (sire unidentified).

A description of the data after having been sub-sampled, according to class of sire, is given in Table 4.3.

The form of the data as outlined was suitable only insofar as the procedure for estimating genetic trend was concerned. On the other hand, the sire evaluation procedure required that records be assembled according to sire, sire-group and a combined herd-year classification so that it was necessary to re-classify the A.I. data as only A.I. sires were evaluated.

Attempts at applying the sire evaluation procedure were unsuccessful because the task was beyond the capacity of the computer (Burroughs B6700). Consequently, the A.I. data was edited and those herd-years or sires represented by less than 5 or 10 records, respectively, were eliminated. A description of the remaining A.I. data (i.e., the edited-A.I. data) is given in Table 4.4.

It is doubtful whether the data collected for this study were representative of the first-lactation records made by the population of Jersey 2-year-olds in the Auckland region during the dairying seasons 1966-67 to 1972-73, inclusive. The requirement that naturally-bred (non-A.I.) cows be sire-identified probably meant that a higher-than-average proportion of pedigree herds was screened for suitable data. Moreover, it is likely that a higher proportion of pedigree herds, relative to all herds, make use of herd-testing services so that a higher-than-average proportion of pedigree animals have recorded lactation performances suitable for study.

Table 4.2 : Statistical description of the complete data.

| Dairying season | Number of lactations | Means and standard errors | | |
|---|---|---|---|---|
| | | Milk yield (lb) | Milkfat yield (lb) | Days in milk |
| 1966-67 | 343 | 5826.8 ± 66.5 | 323.9 ± 3.8 | 280.0 ± 1.8 |
| 1967-68 | 722 | 4833.6 ± 43.3 | 260.0 ± 2.5 | 245.8 ± 1.6 |
| 1968-69 | 1399 | 5515.7 ± 32.7 | 304.1 ± 1.8 | 279.2 ± 0.8 |
| 1969-70 | 1737 | 4638.0 ± 26.0 | 245.5 ± 1.4 | 235.3 ± 0.9 |
| 1970-71 | 1558 | 4752.3 ± 27.9 | 256.8 ± 1.6 | 257.0 ± 1.0 |
| 1971-72 | 1526 | 5151.4 ± 30.1 | 279.1 ± 1.6 | 272.2 ± 0.9 |
| 1972-73 | 1399 | 4934.2 ± 30.6 | 264.5 ± 1.6 | 244.3 ± 1.0 |
| Overall | 8684 | 5001.1 ± 12.8 | 270.3 ± 0.7 | 256.8 ± 0.4 |

Table 4.3 : Description of the data when classified according to sire code.

| Sire class | Number of lactations | Means and standard errors | | |
|---|---|---|---|---|
| | | Milk (lb) | Milkfat (lb) | Days in milk |
| A.I.<br>(111 sires) | 3101 | 5309.7 ± 21.3 | 287.6 ± 1.2 | 259.9 ± 0.7 |
| Non-A.I.<br>(792 sires) | 3869 | 4793.7 ± 18.4 | 259.9 ± 1.0 | 256.4 ± 0.6 |
| Uncoded | 1706 | 4913.0 ± 29.5 | 262.3 ± 1.6 | 252.6 ± 1.07 |

Table 4.4 : Description of the edited-A.I. data - means and standard errors.

| | |
|---|---|
| No. records | 2155 |
| No. sires | 47 |
| No. sire-groups | 6 |
| No. herd-years | 145 |
| Milk yield (lb) | 5238.7 ± 25.2 |
| Milkfat yield (lb) | 282.4 ± 1.4 |
| Days in milk | 258.1 ± 0.8 |

# CHAPTER 5

# RESULTS

## I. ESTIMATES OF GENETIC TREND

### A. Estimates Based on Intra-sire Regression

In Chapter 3, it was stated that the data in this study were sub-sampled according to class of sire; that is, whether the lactation records were produced by the daughters of A.I., non-A.I., or unidentified sires. The estimates of genetic trend in the upper section of Table 5.1 were obtained from the application of the weighted and unweighted intra-sire regression methods of estimation to various combinations of the sub-samples of the data. Mention was also made in Chapter 3 of the need to edit the A.I. data, thereby forming the edited-A.I. data, to allow the sire evaluation procedure to be applied. The estimates of genetic trend in the lower section of Table 5.1 correspond to those in the upper section in that, with the exception of the non-A.I. data, they were obtained from the application of the intra-sire regression analyses to the same combinations of the sub-samples of the data except for the fact that the edited-A.I. data were substituted for the A.I. data.

Table 5.1 : Estimates of genetic trend derived from the application of the intra-sire regression methods of estimation.

| Data | Genetic trend/year (lb) and standard error | | | | Records[(a)] | Reference sires[(b)] | Cross-products[(c)] |
|---|---|---|---|---|---|---|---|
| | Unweighted regression | | Weighted regression | | | | |
| | Milk | Milkfat | Milk | Milkfat | | | |
| | | Analyses involving A.I. data | | | | | |
| A.I. | -79.6 ± 110.6 | -2.3 ± 5.9 | 188.9 ± 39.6 | -3.4 ± 4.2 | 1637 | 67 | 206 |
| Non-A.I. | -89.6 ± 59.0 | -2.0 ± 3.3 | -333.2 ± 26.6 | -16.9 ± 1.5 | 3034 | 81 | 297 |
| A.I., non-A.I. | -104.3 ± 50.7 | -3.3 ± 2.7 | 197.6 ± 25.1 | 8.9 ± 1.4 | 5575 | 155 | 533 |
| A.I., non-A.I., unidentified | -122.0 ± 49.2 | -4.6 ± 2.6 | 126.9 ± 23.3 | 2.9 ± 1.3 | 6875 | 157 | 540 |
| | | Analyses involving edited-A.I. data | | | | | |
| Edited-A.I. | -14.2 ± 137.9 | 2.5 ± 7.2 | 428.4 ± 88.7 | -1.8 ± 4.8 | 1173 | 43 | 131 |
| Edited-A.I., non-A.I. | -79.0 ± 54.8 | -1.0 ± 2.9 | 67.0 ± 26.7 | -0.2 ± 1.45 | 4900 | 128 | 445 |
| Edited-A.I., non-A.I., unidentified | -93.3 ± 53.8 | -2.0 ± 2.8 | 93.8 ± 24.7 | -1.8 ± 1.3 | 5899 | 128 | 448 |

(a) Number of records which contributed to either a daughter (sire) or a contemporary average.

(b) Number of sires (not necessarily different) over which the regression was pooled.

(c) Number of within-sire, within-herd, within-year deviations in the analysis.

Considering the upper section of Table 5.1, the sign and also the consistency of the estimates of genetic trend obtained from the application of the unweighted intra-sire regression to the various sets of data are notable. Overall, these estimates indicate a negative genetic trend of approximately 4 lb. of milkfat and 100 lb. of milk per year. The standard errors indicate that, of the estimates obtained for milkfat yield, the only one that is significantly different from zero is that obtained from the combined set of A.I., non-A.I. and unidentified (all the data in the study), whereas in the case of milk yield, the only estimate not significantly different from zero is that obtained for the A.I. set of data.

The estimates of genetic trend in the upper section of Table 5.1 obtained from the application of the weighted intra-sire regression demonstrate poor agreement with the corresponding estimates obtained using the unweighted regression, particularly as far as milk yield is concerned. Many of the estimates obtained using the weighted regression, are opposite in sign to their counterparts obtained from application of the unweighted regression and, in addition, there is less consistency amongst the estimates obtained from applying the weighted regression to the various sets of data. Without exception, the standard errors of the genetic trend estimates produced by the use of the weighted regression are smaller than the standard errors of the corresponding estimates derived from the unweighted regression analyses. In addition, the standard errors of the estimates produced by the weighted regression are such that only one estimate, the -3.4lb. of milkfat for the A.I. set of data, is not significantly different from zero.

With regard to the estimates of the genetic trend in the lower section of Table 5.1, those obtained from the use of the unweighted regression, with the exception of the edited-A.I. set of data, agree quite well with the analogous estimates of the upper section of the table. None of the milkfat estimates determined by unweighted regression and presented in the lower section of the table is significantly different from zero. The milk yield estimate of the edited A.I. set of data is likewise not significant. Once again, the estimates obtained using the weighted intra-sire regression are in poor agreement with those produced by application of the unweighted regression and in all examples, the standard errors of the estimates from the weighted regression are smaller.

There is no obvious reason for the lack of agreement between the estimates of genetic trend produced from application of the two intra-sire regression methods of estimation, nor is there any real evidence for preferring the estimates of one method over those of the other. Perhaps a point in favour of the unweighted intra-sire regression method of estimation is the consistency of the estimates it produced in this study. However, if it were not for the estimates of -333.2lb. and 428.4 lb. of milk for the non-A.I. and edited-A.I. sets of data, respectively, the estimates obtained using the weighted intra-sire regression are almost as consistent as those obtained using the unweighted regression. Nevertheless, genetic trend estimates of -333.2 lb. and 428.4 lb. of milk per year are difficult to accept and, in fact, the author was prompted to verify by hand the estimate of 428.4 lb. of milk to make sure that the computer programme for applying the weighted intra-sire regression analysis was working correctly. The standard errors of most of the estimates of genetic trend in Table 5.1 make the point that there were insufficient data included in this study to permit reliable estimation of the genetic trend in the dairy cow population from which the lactation records were sampled.

Comparing the number of records which contributed to each estimate of genetic trend in Table 5.1 with the corresponding number available (Tables 4.3 and 4.4), reveals that in terms of utilization of the available data, the intra-sire regression approach for estimating genetic trend is wasteful. It is interesting to note that the degree of utilization of the data is higher when the records of non-A.I. bred animals are involved in the analysis than when they are not; the genetic trend estimate for the A.I. data was based on 53 percent of the available data, whereas 78 percent of the available data contributed to the genetic trend estimate for the non-A.I. set of data. The higher degree of utilization of the data when the records of non-A.I. sires are involved in the regression analyses, stems from the tendency for sires in natural mating (non-A.I. sires), to be used within a herd over a longer time span than A.I. sires and therefore, are generally more able to be represented within a herd by daughter deviations from a contemporary average for the required minimum of three years. The ratios of number of cross-products to reference sires in Table 5.1 substantiate the suggestion of longer usage of non-A.I. sires; the ratio for the non-A.I. data is 3.66 in comparison to 3.07 for the A.I. set of data.

## B. Estimates Based on Mixed Model Solution

In Chapter 3, it was noted that the regression of the mixed model solution of the group effects of the sire evaluation model on time is an estimate of the genetic trend in the population. (Sires are grouped according to the time (year) of entering service and therefore, a group effect is an unweighted genetic average of the sires belonging to that sub-population.) In addition, the regression on time of group effects, weighted by progeny numbers, provides an estimate of the genetic trend attributable to the progeny of the sires. It will be recalled that the MMS method of sire evaluation was applied using two linear models; one in which herd-year effects were treated as fixed effects whereas, in the other, they were regarded as being random effects. The estimates of genetic trend that were obtained from the application of the MMS method of sire evaluation to the edited-A.I. data are given in Table 5.2.

The estimates in Table 5.2 indicate that weighting MMS group effects by the numbers of progeny of individual sires has decreased the extent of the negative genetic trend. Likewise, regarding the herd-year effects in the underlying model to be fixed, instead of random, effects has also tended to reduce the absolute size of the genetic trend estimates. Unfortunately, it was not possible to obtain standard errors of the estimates given in Table 5.2 so it is unknown

Table 5.2 : Estimates of genetic trend produced from the application of the mixed model solution method of estimation to the edited-A.I. data.

| Herd-year effects | Genetic trend/year (lb) | | | |
|---|---|---|---|---|
| | Unweighted | | Weighted | |
| | Milk | Milkfat | Milk | Milkfat |
| Random effects | -76.9 | -3.1 | -69.1 | -2.1 |
| Fixed effects | -66.6 | -2.4 | -61.5 | -2.2 |

whether these estimates are significantly different from zero. In addition, no guidance is provided on how real the changes in the magnitude of the estimates of genetic trend due to the weighting or to the different categories of herd-year effects might be.

In the application of the MMS method of estimation, lactation records from only those herd-years where sire comparisons are possible, contribute to the estimates, i.e., if a herd-year is represented by lactation records produced by the daughters of just a single sire, it is not possible for sire comparisons to occur within that herd-year so that the records concerned have no influence in the analysis. As far as the edited-A.I. data were concerned, every herd-year was represented by lactation records from daughters of at least two sires and therefore, all 2,155 records of the edited-A.I. data contributed to the results given in Table 5.2. In contrast, the estimates of the genetic trend for the edited-A.I. data obtained, using the intra-sire regression analyses, were based on only 1,173 records (Table 5.1). The large difference between the number of records which contributed to the estimates of genetic trend for the edited-A.I. data given in Tables 5.1 and 5.2 precludes these estimates from being the basis of a definitive comparison of the intra-sire regression and MMS procedures for estimating genetic trend.

It is notable that the estimates of genetic trend given in Table 5.2 (MMS) agree quite well overall with those obtained by the application of the unweighted intra-sire regression set out in Table 5.1. In contrast, the estimates in Table 5.1 obtained using the weighted intra-sire regression, particularly with respect to milk yield, show poor agreement with the estimates given in Table 5.2 (MMS).

The lack of agreement amongst the estimates of genetic trend obtained using the various estimation procedures, coupled with the large standard errors of a number of the estimates, prohibits a reliable conclusion being drawn as to the nature of the genetic trend in the data.

## II. VARIANCE COMPONENTS

The estimates of the variance components obtained from applying Henderson's Method I to the edited-A.I. data assuming a fully random model (model C, Chapter 3), together with the corresponding degrees of freedom, are given in Table 5.3. In addition, each variance component is recorded as a percentage of the total variance.

Estimates of variance components have been reported by the Farm Production Division of the N.Z. Dairy Board (N.Z. Dairy Board, 1969). In that study, Henderson's Method I was applied to a sample (unknown number) of age-corrected (mature equivalent) first-lactation records made in the 1967-68 dairying season assuming a model which included herd, sire, and herd-by-sire interaction effects. The results from the study are given in Table 5.4 (again each variance component is recorded as a percentage of the total variance).

The results in Tables 5.3 and 5.4 show that the absolute values of the estimates of the variance components common to both studies are markedly different. This could be attributable to the lactation records analyzed in the N.Z. Dairy Board's study having been age corrected (the correction factors were partly multiplicative) whereas, the records were not age corrected in the present study. Evidence on this point is provided by the estimates of variance components obtained in a small-scale study by Searle (1961b). In his study, Henderson's Method I was applied to a sample of 591 uncorrected first-lactation milkfat records produced in 88 herds in the 1954-55 dairying season assuming the same model as that in the study reported by the N.Z. Dairy Board. The results of Searle's study are given in Table 5.5.

Table 5.3 : Variance component estimates obtained from the edited-A.I. data. The components as a percentage of the total variance are given in parentheses.

| Effect | | Estimate | | Degrees of freedom |
|---|---|---|---|---|
| | | Milk | Milkfat | |
| Sire group | $\hat{\sigma}^2_g$ | 3,339 (0.2) | 25 (0.6) | 5 |
| Sires within group | $\hat{\sigma}^2_s$ | 114,367 (8.3) | 273 (6.5) | 41 |
| Herd-year | $\hat{\sigma}^2_h$ | 386,977 (28.0) | 1,425 (33.9) | 144 |
| Herd-year x sire within group interaction | $\hat{\sigma}^2_{hs}$ | -61,453 (-4.4) | -239 (-5.7) | 562 |
| Error | $\hat{\sigma}^2_e$ | 939,612 (67.9) | 2,721 (64.7) | 1,062 |
| Total variance | | 1,382,842 | 4,205 | |

Table 5. 4 : Variance component estimates reported by the N. Z. Dairy Board. The components as a percentage of the total variance are given in parentheses.

| Effect | Estimate | |
|---|---|---|
| | Milk | Milkfat |
| Herd | 819,000 (33.5) | 2,722 (36.8) |
| Sire | 228,000 (9.3) | 442 (6.0) |
| Herd x sire interaction | -22,000 (-0.9) | 98 (1.3) |
| Error | 1,419,000 (58.1) | 4,132 (55.9) |
| Total variance | 2,444,000 | 7,394 |

Table 5.5 : Variance component estimates (milkfat only) reported by Searle (1961b).

| Effect | Estimate | % of total |
|---|---|---|
| Herd | 2,040 | 48 |
| Sire | 80 | 2 |
| Herd-by-sire interaction | 50 | 1 |
| Error | 2,040 | 49 |
| Total | 4,210 | |

The common ground for making comparisons of the absolute values of the estimates of the variance components obtained from the different studies (Tables 5.3, 5.4 and 5.5) are the respective estimates of the error variance component since, in all the studies, the error variance represents within-herd residual variation. As Searle's study did not involve milk records, the estimates in the three studies may only be compared using estimates of variance components for milkfat yield. The results in Tables 5.3, 5.4 and 5.5 demonstrate that whilst the estimates of the error variance component produced by analyzing uncorrected records, that is, the present study (2,721) and Searle's study (2,040) are similar. Both, however, are markedly different from the estimate produced by the N.Z. Dairy Board's analysis of age-corrected records (4,132). Thus, it appears that age corrections were largely responsible for the previously noted disagreement between the size of the estimates obtained in the present study and the size of the estimates reported by the N.Z. Dairy Board.

For comparative purposes, the problem of variation in the absolute size of estimates of variance components obtained from different studies may be overcome by expressing the estimates of individual components as percentages of the estimate of the total variance of the records contributing to the respective analyses. The results in Tables 5.3 and 5.4 demonstrate that in the present study, and that reported by the N.Z. Dairy Board, variance between sires has accounted for virtually the same proportion of the total variation. The other common variance component in these studies is the error variance, but in this case, the agreement between the two analyses is not as good. Most likely, the lack of agreement between the proportions of the total variation attributable to error sources is a result of sampling variation in the data included in the two studies. Likewise, sampling variation probably accounts for the lack of agreement between the percentage estimates obtained in Searle's study (Table 5.5) and those obtained in the other studies (Tables 5.3 and 5.4).

The reason for estimating variance components in the present study was to obtain estimates of the ratios:

$$\frac{\sigma_e^2}{\sigma_s^2} \quad \text{and} \quad \frac{\sigma_e^2}{\sigma_h^2} .$$

However, it is informative to use the appropriate estimates of variance components to obtain an estimate of the within-herd heritability based on the expression: $4\,\sigma_s^2 \,/\, (\sigma_s^2 + \sigma_e^2)$. The estimated ratios of variance components obtained from the present as well as the N. Z. Dairy Board analyses, together with heritability estimates, are given in Table 5.6.

The notable feature of the results in Table 5.6 is the similarity, especially as far as milkfat yield is concerned, between the estimates.

In the study reported by the N. Z. Dairy Board, the model employed can be considered to be a fully random model and therefore, the application of Henderson's Method I is expected to produce unbiased estimates of the variance components. If this is so, the results given in Table 5.6 suggest that the decision to regard the fixed effects (group effects) of the model in the present study as being random effects (to make Henderson's Method I applicable), may not have biased the estimates of the variance components of the random effects in the model to any great extent.

## III. ESTIMATES OF SIRE BREEDING VALUES

The MMS method of sire evaluation was applied to the edited-A.I. data using two linear models; one in which herd-year effects were regarded as being random effects, i.e., model A (Chapter 3), the other, a variant of model A whereby herd-year effects were considered to be fixed effect. The estimates of the sire breeding

Table 5.6 : Estimates of variance component ratios and heritabilities.

| Trait | Study | Estimate [(a)] | | |
|---|---|---|---|---|
| | | $\hat{\sigma}_e^2/\hat{\sigma}_s^2$ | $\hat{\sigma}_e^2/\hat{\sigma}_h^2$ | Heritability |
| Milk yield | Present | 8.22 | 2.43 | 0.43 |
| | N.Z. Dairy Board | 6.22 | - | 0.55 |
| Milkfat yield | Present | 9.98 | 1.91 | 0.36 |
| | N.Z. Dairy Board | 9.35 | - | 0.39 |

(a) Based on the estimates given in Tables 5.3 and 5.4.

values for milk and milkfat yield obtained from these analyses, together with the corresponding N. Z. Dairy Board milkfat sire ratings (as at the 1973 mating season), are given in Table 5. 7. (As a number of sires did not have a N. Z. Dairy Board rating for milk yield, none is given in the table.) Under the sire evaluation model used in the present study, the sires were grouped on the basis of year of entering service and in Table 5. 7 sires have been spaced accordingly.

The important feature of the results given in Table 5. 7 is that for many sires, the MMS estimates of their breeding values were obtained from a very small number of daughter records. Therefore, the estimates given in Table 5. 7 do not constitute a satisfactory basis for definitive comparison of the MMS and N. Z. Dairy Board methods of sire evaluation. Nevertheless, Spearman rank correlation analyses were applied to the estimates in Table 5. 7 in order to obtain an idea of the degree of relationship between the different methods of sire evaluation. The rank correlations amongst the estimates of Table 5. 7 are given in Table 5. 8.

The Spearman rank correlation coefficients were calculated in recognition of the point that the estimates of breeding value given in Table 5. 7 constitute a selected sample and therefore, product-moment correlation analyses are not appropriate. The rank correlation estimates given in Table 5. 8 demonstrate that in applying the MMS method of evaluation in this study, the use of a model in which herd-year effects are regarded as being fixed, as opposed to random effects, would be of little consequence as far as the ranking of the sires involved is concerned. In contrast, the degree of relationship between the estimates of breeding values obtained by the MMS and N. Z. Dairy Board methods of evaluation is very low. However, it has already been noted that factors apart from actual differences between the MMS and N. Z. Dairy Board methods of evaluation are likely to have played a major role in bringing about the low correlation between the estimates obtained from the two methods in this case. Thus, the low correlation between the MMS and N. Z. Dairy Board estimates of breeding value given in Table 5. 7 are

Table 5.7 : Estimates of sire breeding values. (Sire rankings are given in parentheses.)

| Sire[a] | Daughters (lactations) | Estimate (lb) | | | | |
|---|---|---|---|---|---|---|
| | | MMS | | | | N. Z. Dairy Board rating[b] |
| | | Herd-year effects | | Herd-year effects | | |
| | | Fixed | Random | Fixed | Random | |
| | | Milk | | Milkfat | | Milkfat |
| 1 | 22 | 6219 (3) | 5684 (3) | 339 (3) | 305 (3) | 30 (27) |
| 2 | 10 | 6220 (2) | 5809 (1) | 338 (4) | 309 (1) | 24 (36) |
| 3 | 9 | 6053 (9) | 5588 (6) | 330 (8) | 299 (6) | 26 (33) |
| 4 | 11 | 5833 (26) | 5308 (24) | 324 (17) | 290 (15) | 20 (41) |
| 5 | 17 | 5834 (25) | 5294 (25) | 330 (8) | 296 (8) | 22 (40) |
| 6 | 11 | 5871 (21) | 5406 (18) | 313 (36) | 280 (30) | 24 (36) |
| 7 | 17 | 5996 (13) | 5569 (7) | 331 (7) | 300 (5) | 24 (36) |
| 8 | 39 | 5972 (14) | 5483 (14) | 330 (8) | 296 (8) | 40 (7) |
| 9 | 26 | 62[illegible] (4) | 5674 (4) | 328 (13) | 292 (13) | 20 (41) |
| 10 | 21 | 601[illegible] ([illegible]) | 5417 (17) | 321 (21) | 283 (26) | 27 (31) |

Table 5.7 Cont'd.

| | | | | | | |
|---|---|---|---|---|---|---|
| 11 | 26 | 6175 (5) | 5647 (5) | 327 (14) | 292 (13) | 35 (18) |
| 12 | 71 | 5716 (31) | 5212 (29) | 318 (25) | 285 (22) | 35 (18) |
| 13 | 36 | 6046 (10) | 5492 (13) | 323 (18) | 287 (19) | 34 (22) |
| 14 | 59 | 5904 (19) | 5428 (16) | 310 (38) | 277 (36) | 26 (33) |
| 15 | 79 | 6089 (6) | 5528 (8) | 334 (5) | 297 (7) | 39 (10) |
| 16 | 11 | 5942 (15) | 5398 (19) | 329 (12) | 293 (12) | 32 (25) |
| 17 | 58 | 6013 (12) | 5499 (11) | 340 (2) | 305 (3) | 49 (2) |
| 18 | 13 | 5713 (32) | 5173 (31) | 315 (31) | 279 (32) | 15 (44) |
| 19 | 37 | 6062 (8) | 5507 (9) | 332 (6) | 294 (10) | 36 (16) |
| 20 | 183 | 5747 (28) | 5225 (28) | 325 (16) | 290 (15) | 37 (13) |
| 21 | 40 | 5837 (24) | 5288 (26) | 311 (37) | 275 (38) | 15 (44) |
| 22 | 20 | 5672 (34) | 5162 (32) | 314 (33) | 280 (30) | 38 (12) |
| 23 | 89 | 6330 (1) | 5807 (2) | 342 (1) | 307 (2) | 49 (2) |
| 24 | 11 | 5933 (16) | 5452 (15) | 322 (19) | 290 (17) | 17 (43) |
| 25 | 15 | 5628 (37) | 5124 (34) | 317 (28) | 285 (22) | 43 (4) |
| 26 | 36 | 5563 (42) | 5039 (40) | 306 (41) | 272 (39) | 26 (33) |
| 27 | 22 | 5930 (17) | 5503 (10) | 318 (27) | 287 (19) | 34 (22) |
| 28 | 27 | 5406 (44) | 4922 (44) | 293 (47) | 259 (46) | 27 (31) |
| 29 | 9 | 5425 (43) | 4927 (43) | 303 (44) | 270 (41) | 15 (44) |
| 30 | 35 | 5818 (27) | 5280 (27) | 318 (25) | 282 (27) | 41 (5) |
| 31 | 20 | 5638 (35) | 5111 (35) | 314 (33) | 279 (32) | 39 (10) |

Table 5.7 Cont'd.

| | | | | | | |
|---|---|---|---|---|---|---|
| 32 | 130 | 5363 (45) | 4806 (46) | 308 (39) | 271 (40) | 40 (7) |
| 33 | 94 | 5682 (33) | 5110 (36) | 319 (24) | 282 (27) | 36 (16) |
| 34 | 15 | 5324 (46) | 4907 (45) | 295 (45) | 264 (45) | 34 (22) |
| 35 | 13 | 5844 (23) | 5321 (23) | 322 (19) | 287 (19) | 32 (25) |
| 36 | 30 | 5719 (30) | 5138 (33) | 304 (43) | 266 (44) | 30 (27) |
| 37 | 150 | 5918 (18) | 5364 (20) | 321 (21) | 284 (24) | 40 (7) |
| 38 | 97 | 5291 (47) | 4719 (47) | 295 (45) | 257 (47) | 10 (47) |
| 39 | 93 | 5626 (38) | 5060 (38) | 321 (21) | 284 (24) | 41 (5) |
| 40 | 110 | 5896 (40) | 5353 (21) | 317 (28) | 281 (29) | 37 (13) |
| 41 | 18 | 6068 (7) | 5495 (12) | 327 (14) | 290 (17) | 30 (27) |
| 42 | 96 | 5745 (29) | 5193 (30) | 315 (31) | 278 (35) | 35 (18) |
| 43 | 80 | 5854 (22) | 5322 (22) | 330 (8) | 294 (10) | 52 (1) |
| 44 | 29 | 5609 (39) | 5048 (39) | 305 (42) | 267 (43) | 24 (36) |
| 45 | 12 | 5595 (40) | 5027 (42) | 307 (40) | 270 (41) | 30 (27) |
| 46 | 68 | 5580 (41) | 5039 (40) | 314 (33) | 277 (36) | 37 (13) |
| 47 | 40 | 5629 (36) | 5080 (37) | 316 (30) | 279 (32) | 35 (18) |

(a) N.Z. Dairy Board sire codes are given in Appendix 1.

(b) Source: N.Z. Dairy Board 1973 : Artificial Breeding Service, Reference Catalogue.

Table 5.8 : Spearman rank correlations amongst the breeding value estimates given in Table 5.7. (Degrees of freedom = 45).

| Method | | MMS | | | |
|---|---|---|---|---|---|
| | | Herd-years fixed | | Herd-years random | |
| | | Milk | Milkfat | Milk | Milkfat |
| N. Z. Dairy Board | Milkfat | | 0.25 N.S. | - | 0.22 N.S. |
| MMS Herd-year random | Milk | 0.98*** | - | | |
| | Milkfat | | 0.98*** | | |

N.S. = $p > 0.05$. *** = $p < 0.001$.

unlikely to reflect the true differences (if they occur) between the two methods as far as sire rankings are concerned.

Sires belonging to the N. Z. Dairy Board's proven bull team in the 1973 mating season all had both a milk and milkfat rating and therefore, it was decided that a comparison of the MMS and N. Z. Dairy Board methods of evaluation based on breeding value estimates for both milk and milkfat yield might be informative. The estimates of the proven sire breeding values for milk and milkfat yield obtained from application of the MMS and N. Z. Dairy Board methods of evaluation are given in Table 5.9.

The results of Spearman rank correlation analyses of the estimates given in Table 5.9 are recorded in Table 5.10.

Although the average number of lactation records which contributed to the MMS estimate of the breeding values of the sires in the N. Z. Dairy Board's proven bull team for 1973 is larger than that when all the sires of this study are included (48 and 62, respectively), a definitive comparison of the MMS and N. Z. Dairy Board methods of evaluation is not possible. The result of particular interest in Table 5.10 is the low correlation between the milk and milkfat ratings derived from the N. Z. Dairy Board's method of sire evaluation, especially when compared to the corresponding correlations for the MMS methods of evaluation. The low correlation between the N. Z. Dairy Board milk and milkfat ratings is probably due to the fact that milk yield ratings have only recently been introduced by the Board and consequently, a sire's milk rating is often based on fewer daughter records than is the case for his milkfat rating. It is also of interest to note that in Table 5.10 the correlations between the N. Z. Dairy Board's milk ratings and the MMS estimates of breeding value for milk yield are higher than the corresponding correlations for milkfat yield. As before, the appropriate correlation estimates in Table 5.10 indicate that in this study, regarding herd-year effects to be either fixed or random effects has had little effect on sire rankings produced by the MMS method of sire evaluation.

Table 5.9 : Estimated breeding values of the sires belonging to the N.Z. Dairy Board's proven bull team in the 1973 mating season. (Sire rankings are given in parentheses.)

| Sire | Daughters (lactations) | Estimated breeding value | | | | | |
|---|---|---|---|---|---|---|---|
| | | Milk | | | Milkfat | | |
| | | MMS | | N.Z. Dairy Board rating[(a)] | MMS | | N.Z. Dairy Board rating[(a)] |
| | | Herd-year fixed | Herd-year random | | Herd-year fixed | Herd-year random | |
| 8 | 39 | 5972 (7) | 5483 (8) | 631 (11) | 330 (5) | 296 (4) | 40 (7) |
| 11 | 26 | 6175 (2) | 5647 (2) | 877 (1) | 327 (7) | 292 (7) | 35 (17) |
| 13 | 36 | 6046 (5) | 5492 (7) | 732 (4) | 323 (9) | 287 (9) | 34 (20) |
| 15 | 79 | 6089 (3) | 5528 (3) | 697 (6) | 334 (3) | 297 (3) | 39 (10) |
| 17 | 58 | 6013 (6) | 5499 (6) | 597 (13) | 340 (2) | 305 (2) | 49 (2) |
| 19 | 37 | 6062 (4) | 5507 (4) | 675 (8) | 332 (4) | 294 (5) | 36 (15) |
| 20 | 183 | 5747 (13) | 5225 (13) | 406 (20) | 325 (8) | 290 (8) | 37 (12) |
| 23 | 89 | 6330 (1) | 5807 (1) | 864 (2) | 342 (1) | 307 (1) | 49 (2) |
| 25 | 15 | 5628 (19) | 5124 (16) | 342 (23) | 317 (15) | 285 (11) | 43 (4) |

**Table 5.9 Cont'd.**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 26 | 36 | 5563 (24) | 5039 (22) | 487 (16) | 306 (24) | 272 (21) | 26 (26) |
| 27 | 22 | 5930 (8) | 5503 (5) | 765 (3) | 318 (13) | 287 (9) | 34 (20) |
| 28 | 27 | 5406 (25) | 4922 (25) | 653 (10) | 293 (28) | 259 (27) | 27 (25) |
| 30 | 35 | 5818 (12) | 5280 (12) | 701 (5) | 318 (13) | 282 (14) | 41 (5) |
| 31 | 20 | 5638 (17) | 5111 (17) | 487 (16) | 314 (19) | 279 (17) | 39 (10) |
| 32 | 130 | 5363 (26) | 4806 (27) | 258 (26) | 308 (21) | 271 (22) | 40 (7) |
| 33 | 94 | 5682 (16) | 5110 (18) | 474 (18) | 319 (12) | 282 (14) | 36 (15) |
| 34 | 15 | 5324 (27) | 4907 (26) | 333 (25) | 295 (26) | 264 (26) | 34 (20) |
| 36 | 30 | 5719 (15) | 5138 (15) | 617 (12) | 304 (25) | 266 (25) | 30 (23) |
| 37 | 150 | 5918 (9) | 5364 (9) | 692 (7) | 321 (10) | 284 (12) | 40 (7) |
| 38 | 97 | 5291 (28) | 4719 (28) | -35 (28) | 295 (26) | 257 (27) | 10 (28) |
| 39 | 93 | 5626 (20) | 5060 (20) | 337 (24) | 321 (10) | 284 (12) | 41 (5) |
| 40 | 110 | 5896 (10) | 5353 (10) | 657 (9) | 317 (15) | 281 (16) | 37 (12) |
| 42 | 96 | 5745 (14) | 5193 (14) | 390 (22) | 315 (18) | 278 (19) | 35 (17) |
| 43 | 80 | 5854 (11) | 5322 (11) | 536 (14) | 330 (5) | 294 (5) | 52 (1) |
| 44 | 29 | 5609 (21) | 5048 (21) | 258 (26) | 305 (23) | 267 (24) | 24 (27) |
| 45 | 12 | 5595 (22) | 5027 (24) | 489 (15) | 307 (22) | 270 (23) | 30 (23) |
| 46 | 68 | 5580 (23) | 5039 (22) | 397 (21) | 314 (19) | 277 (20) | 37 (12) |
| 47 | 40 | 5629 (18) | 5080 (19) | 414 (19) | 316 (17) | 279 (17) | 35 (17) |

(a) Source: N.Z. Dairy Board's Artificial Breeding Service, Reference Catalogue, 1973.

Table 5.10 : Spearman rank correlation coefficients based on the estimates of breeding value for the sires of the N.Z. Dairy Board's proven bull team in the 1973 mating season. (Degrees of freedom = 26).

| Method | | | MMS | | | | N.Z. Dairy Board | |
|---|---|---|---|---|---|---|---|---|
| | | | Herd-years fixed | | Herd-years random | | | |
| | | | Milk | Milkfat | Milk | Milkfat | Milk | Milkfat |
| MMS | Herd-years fixed | Milk | | 0.86*** | 0.99*** | | 0.81*** | |
| | | Milkfat | | | | 0.98*** | | 0.69*** |
| | Herd-years random | Milk | | | | 0.88*** | 0.80*** | |
| | | Milkfat | | | | | | 0.67*** |
| N.Z. Dairy Board | | Milk | | | | | | 0.17 N.S. |
| | | Milkfat | | | | | | |

# CHAPTER 6

# DISCUSSION

In considering the extent to which the results of the present study can be generalized, it is necessary to take into account some aspects of the sample of data which was analyzed. In Chapter 4, it was indicated that because of limited resources, the lactation records of only those 2-year-old heifers that were of the Jersey breed and milked in the area serviced by the Auckland Herd Improvement Association, were screened for data. Furthermore, it was noted that in meeting the requirements of the estimation procedures employed in the study, it was probable that a higher-than-average proportion of lactation records produced by pedigree animals was sampled. The figures in Table 4.3 show that the data included lactation records from a larger number of non-A.I. bred animals than A.I.-bred animals (3,869 and 3,101, respectively) which suggests that the extent of A.I. usage (i.e., proven sire usage) in the herds screened for data was not large. These characteristics of the data need to be borne in mind in drawing conclusions from the study.

An important attribute of the data was that all were first-lactation records. In this connection, Powell, Spike and Meadows (1973), in a study of the characteristics of first-lactation records commented: "The first lactation should be the most meaningful, generally free from confusions of dry period, previous production, selection [culling], and cumulative disorders not necessarily attributable to a cow's [genetic] merit." Moreover, as the records included in the present study were produced by cows all of the same age (2-year-olds), corrections for age were unnecessary and consequently, the difficulties associated with their use (considered in Chapter 2), were avoided.

Another aspect of the data requiring comment is that no general adjustment of the lactation records for variation in days in milk (lactation length) was made, other than exclusion of records from lactations shorter than 100 days and correction of records made in an interval exceeding 305 days. Dealing with variation in lactation length in this manner is not expected to have biased the results of the present study to any great extent. Van Vleck (1962) found that the bias in sire evaluation from ignoring incomplete first-lactation records (i.e., produced in less than 305 days; terminated by voluntary culling, but not including the records of cows genuinely dry), was small and unimportant. Powell <u>et al</u>. (1973) showed that although the method of extending voluntarily terminated first-lactation records (less than 305 days) had a significantly different effect among sires (significant sire-by-method of extension interaction; $p < 0.05$), it had little effect on sire ranking.

A New Zealand study (N.Z. Dairy Board, 1958), demonstrated that 3.2 percent of first lactations were terminated in less than 100 days because of low production, failure to milk, disease or accident. In the same study, an analysis of variance showed that variation in the lactation length of 2-year-old heifers was significantly different between sires ($p < 0.01$) and that the heritability of first-lactation length was 0.16. These results drew the comment: "By excluding lactations of less than 100 days some bulls will be overrated, but from these figures it is likely that the daughters remaining in the [sire] survey will also be below average in lactation length." However, in a later and more extensive study (N.Z. Dairy Board, 1969), the heritability of first-lactation length was estimated to be 0.09. This suggests that the absence of routine correction for variation in lactation length is unlikely to bias the results of studies utilizating only first-lactation records.

In Chapter 5, it was commented that disagreement amongst the estimates of genetic trend, allied with generally large standard errors, precluded reliable conclusions as to the nature of the genetic trend in the data. Nevertheless, the point of interest, particularly with respect to those estimates obtained by application of the unweighted intra-sire regression and MMS methods of estimation, is that, despite expectations (based on studies reported in the literature; considered in Chapter 2), it is unlikely that the genetic trend in the data of the present study was positive. In view of surveys conducted by the Farm Production Division of the N.Z. Dairy Board, this is quite an interesting result. For instance, from knowledge of sire ratings and sire usage patterns, it has been estimated (N.Z. Dairy Board, 1974) that the use of proven sires in A.I. in the period bounded by the 1956-57 and 1972-73 dairying seasons, has raised the production of the national herd and herds having continuously used A.I. by 19.3 lb. and 42.7 lb. of milkfat, respectively. The corresponding estimates for the period covered in the present study (1966-67 and 1972-73 dairying seasons, inclusive) are 13.4 lb. and 21.0 lb. of milkfat representing an annual genetic gain of 1.91 lb. and 3.0 lb. of milkfat, respectively. The results of the present study generally do not concur with its estimates.

The fact that the estimates of genetic trend reported by the N.Z. Dairy Board (1974) are predictions based on estimates of progeny superiority rather than estimates of the actual genetic change in the population, may partly explain their lack of agreement with the results of the present study. It was noted in Chapter 2 that seldom have the estimates of actual genetic change in dairy cattle populations reached the levels predicted on the basis of estimates of progeny superiority. Furthermore, it was shown that in an improving population, the delay between conception and completion of a first lactation by the heifer means that, relative to actual genetic trend in the population, the prediction approach is effectively four years "out-of-date". In a discussion on relating progeny superiority to genetic trend in cattle, P. D. Miller (1969) commented: "Genetic change, however, depends not on the superiority of progeny to the parent population but on the superiority of replacements to the existing population, the proportion of the population which is replaced, and the culling within the existing population." Clearly, genetic trend is a function of not only the genetic superiority of sires available to a dairy cow population through A.I., but also the selection and breeding practices within individual herds. In this context, the fact that it is likely that a higher than average number of pedigree herds was screened for data in the present study, has important implications.

It is well known, that farmers owning pedigree (registered) herds do not rigidly adhere to the principles of selection based on productive performance and that "type" characteristics, of doubtful importance, are often given undue emphasis. Thus, the absence of a positive genetic trend in the data may reflect a lack of selection pressure for milk and milkfat production in the herds screened for the data included here. It has already been noted that a larger number of non-A.I. bred animals than A.I.-bred animals were represented in the data (cf. Table 4.3) and therefore, the lack of positive genetic trend in the present study may also have been due to insufficient use of sires reliably evaluated as being of superior genetic merit. Considering these points, the estimates of genetic trend obtained most likely reflect the special characteristics of the data and not the genetic trend in the overall dairy cow population in this country.

A result of particular interest which also relates to the estimates of genetic trend, is the very poor agreement (frequently involving opposite sign) between the estimates obtained from application of the weighted, as opposed to the unweighted intra-sire regression methods of estimation (Table 5.1). Although by no means conclusive, there are a few points which suggest that possibly it is the estimates obtained by the application of the weighted intra-sire regression method of estimation that could be at fault. These are:

(i) The estimates of genetic trend obtained using the MMS method of estimation agree much more closely with the estimates obtained using the unweighted, as opposed to the weighted intra-sire regression method.

(ii) When compared across the different sets of data, the estimates of genetic trend produced by applying the weighted regression (Table 5.1) are less consistent than those obtained using the unweighted intra-sire regression method.

(iii) The estimates of genetic trend per year for milk yield of -333.3 lb. (non-A.I. data) and 428.4 lb. (edited-A.I. data) obtained from applying the weighted intra-sire regression are extraordinary and difficult to accept.

If the estimates obtained using the weighted intra-sire regression method are, in fact, in error, there appears to be two main areas where this may have arisen. These are:

(i) Incorrect computation of the estimates.

(ii) The method may be inappropriate from a theoretical standpoint.

Incorrect computation cannot be unequivocally rejected, but it does seem to be unlikely since the computer programme which was written to apply the weighted and unweighted intra-sire regression methods of estimation was tested in a number of sample runs prior to being applied to the actual data. However, a point about this approach in establishing accurate computer programmes is that in devising test analyses, the programmer believes that he is able to predict all the peculiarities which may arise and which are capable of leading the programme into producing incorrect results. Under certain circumstances, particularly when unbalanced data are being analyzed, this belief may be presumptuous. The fact that the genetic trend estimate for milk yield in the edited-A.I. data (428.4 lb/year) was verified exactly by hand calculation suggests that computational errors may not be the explanation for the questioned results.

It is doubtful that the weighted intra-sire regression method for estimating genetic trend is inappropriate from a theoretical standpoint. According to Hickman (1971), the theoretical expectation of the regression coefficient concerned, indicated that the weighted intra-sire regression method does yield an unbiased estimate of the genetic trend in a population. In addition to having been successfully applied (Hickman, 1971), the weighted intra-sire regression technique for estimating genetic trend represents a logical sequence of ideas.

The idea of weighting a within-herd and within-year production deviation of A.I.-bred, as opposed to naturally-bred animals by the inverse of the variance of the deviation as a way of increasing the reliability of the comparison, was introduced by Robertson and Rendel (1954). In order that such comparisons be unbiased by any genetic trend in the naturally-bred population, Van Vleck and Henderson (1961b) proposed a least squares procedure which was based on the analysis of weighted within-herd and within-year production deviations (first-lactation records only) of the A.I.- and naturally-bred populations. Smith (1962) drew attention to the fact that the regression of within-sire, within-herd, and within-year production deviations on time represents minus one-half the overall genetic trend in the population. If the deviations in the method proposed by Van Vleck and Henderson (1961b) are calculated on a within-sire as well as a within-herd and within-year basis, and furthermore, if no weighting factors are used, their method is equivalent to that outlined by Smith (1962). On the other hand, Hickman (1971) proposed that within-sire, within-herd, and within-year deviations calculated in applying the method proposed by Smith (1962) be weighted by the inverse of their variances before being regressed on time. If the method proposed by Van Vleck and Henderson (1961b) is applied on a within-sire as well as a within-herd and within-year subclass basis, the resulting technique is equivalent to the method suggested by Hickman (1971); i.e., the weighted intra-sire regression method of estimation used in the present study.

The other aspect in the theory of the weighted intra-sire regression approach relates to the weighting factor itself. In the present study, weighting factors of size 10 (or more) were regularly encountered so that a deviation with a corresponding weighting factor of 10 carries ten times the "weight" of a deviation with a weighting factor of one. It could be, that under certain circumstances, weighting a within-herd production deviation by the inverse of its variance tends to "over emphasize" the reliability of the deviation.

If this were the case, the effect is expected to be more important when only a small set of data is being analyzed; for instance, in the edited-A.I. data.

Unfortunately, the disagreement between the estimates of genetic trend produced by the two intra-sire regression techniques remains unexplained.

The way in which non-random mating effects the estimates of genetic trend produced by applying intra-sire regression techniques was considered in Chapter 2. For instance, it was noted that if older sires tend to be mated to older cows, estimates of genetic trend which are based on change in performance of half sisters over time (e.g., intra-sire regression methods), are downwardly biased. However, Smith (1962) noted that the bias is avoided if the intra-sire regression techniques are applied on a within-age of dam basis. While this is acceptable from a theoretical standpoint, in practice the suggestion may be unworkable. In Chapter 5, attention was drawn to the high rate of data rejection when intra-sire regression methods are used and, therefore, if yet another classification is introduced (i.e., deviations calculated on a within-age of dam, within-sire, within-herd, and within-year basis), the problem of the low rate of utilization of the input data would be aggravated. Thus, where non-random mating is thought to be important, a more realistic approach in overcoming the bias may be that adopted by Syrstad (1966) of estimating the magnitude of the bias and correcting the estimate of genetic trend. In the present study, age of dam was not recorded so that is is not known whether non-random mating may have affected the estimates of genetic trend produced by application of the intra-sire regression methods of estimation.

In the past, the intra-sire regression methods of estimating genetic trend have had an advantage over the MMS method of estimation (and variations thereof) in that they place less demand on computing facilities. However, the development of modern computers has probably removed that advantage. Based on the comments of Henderson (1973), it would seem that skilfull programming and modern computers have made it possible for the MMS method of estimating genetic trend to be applied to almost unlimited data. In the present study, a notable advantage of the MMS method for estimating genetic trend, in comparison with the intra-sire regression techniques, was its higher degree of utilization of input data. Had the standard errors of the estimates of genetic trend produced by applying the MMS method of estimation been obtained, it is likely that a comparison with the standard errors of the genetic trend estimates produced by the intra-sire regression methods would have reflected the difference in utilization of input data of the two approaches.

At one stage it was thought that the extraordinary estimates of genetic trend obtained from applying the weighted intra-sire regression (-333.2 lb. and 428.4 lb of milk per year) may have been a result of a peculiarity of the data. However, if there had been something unusual about the data, it would almost certainly have been reflected in the estimates of variance components, particularly the estimate of the error variance component obtained from the data. The results in Chapter 5, Section II demonstrate that the estimates of variance components obtained are in good agreement with those previously reported (N.Z. Dairy Board, 1969) and, therefore, any suggestion of peculiarity in the data of the present study is unlikely. In this connection, the results in Table 5.6 are of particular interest in that the estimates of heritability (and therefore, the estimates of the error-to-sire variance ratios) obtained agree rather well with those reported by the N.Z. Dairy Board (1969). Moreoever, the estimates of heritability in Table 5.6 are not greatly different from estimates of the heritability of first-lactation traits that have been reported elsewhere.

In an investigation of 2,720 first-lactation records produced in 1947 by heifers of a number of breeds, Rendel, Robertson, Asker, Khishin and Ragab (1957) obtained an estimate of heritability for milk yield of 0.50 ± 0.08. (Heritability estimate based on intra-class correlation amongst half-sibs performance.) Van Vleck, Wadell and Henderson (1961) applied Henderson's Method I, assuming a model which included sire, year-season, herd, and all possible interaction effects to a sample of 39,728 first-lactation records. The records were produced in 3,932 herds and 27 year-seasons and were age corrected to mature equivalent before analysis. Application of the technique used to calculate heritability in the present case to the estimates of variance components published by Van Vleck <u>et al.</u>, produced heritability estimates of 0.43 and 0.48 for milk and milkfat yield, respectively.

That the method of age adjustment and level of herd production may bias estimates of heritability has been demonstrated by Norman, McDaniel and Dickinson (1972). They studied 7,997 records made by cows calving for the first time in November and December 1964 at the age of 26 to 28 months. (The strict limitations on the data were aimed at minimizing the effect of any inter-relationship between age and month of calving on milk yield, cf. P.D. Miller <u>et al.</u>, 1970.) Regional age-correction factors were used to correct to both a mature-equivalent level and to a 27 month-equivalent level. The heritability estimates that were obtained from analyzing the mature-equivalent data after pooling across the regions, were 0.66 and 0.62 for milk and milkfat yield, respectively. The corresponding estimates for the 27 month-equivalent records were 0.50 and 0.46 which clearly indicates that the age adjustment base has an effect on heritability estimates. The heritability estimates that were obtained from pooling within-region estimates of variance components (weighted by number of observations) for the mature-equivalent data were 0.49 and 0.45 for milk and milkfat, respectively. The corresponding estimates for the 27 month-equivalent data were 0.52 and 0.46; that is, no change in the case of milkfat yield and a change of only 0.02 in the case of milk yield. The latter differences are much less than those for the mature-equivalent data. The results obtained by Norman <u>et al.</u> (1972) indicate that the use of regional age-correction factors having a base age quite different from the age at calving of the cows which produced the records analyzed, can seriously bias heritability estimates. In this respect it is worth noting that the heritability estimates reported by the N.Z. Dairy Board (1969) were obtained by pooling within-region estimates of variance components from mature-equivalent data. (Regions represented the two areas serviced by the Awahuri and Newstead A.I. centres.) Norman et al. (1972) also demonstrated that stratifying the data according to level of herdmate production and later pooling the variance component estimates, leads to different estimates of heritability than when level of production is ignored. Thus, it appears that close attention should be paid to choosing an appropriate model on which to base heritability estimates.

The heritability estimates reported by the N.Z. Dairy Board (1969), Van Vleck et al. (1961), Norman et al. (1972), as well as those of the present study, were based on the records of only A.I.-bred animals and, therefore, may be biased through the A.I. sires being a selected group. (The daughters of a selected group of sires will not have arisen through random mating in the population and consequently, the between sire variance for the population, and the heritability, may be underestimated.) Furthermore, sire effects are likely to be correlated with the number of observations (daughters) in sire subclasses which is also a source of bias in estimates of variance components (Harville, 1968; Henderson, 1973). The importance of these biases in the estimates obtained in the present study is not known.

Despite being a main objective of the present study, it has not been possible to determine whether the method of sire evaluation used by the Farm Production Division of the N.Z. Dairy Board (N.Z. Dairy Board, 1970) ranks sires (based on estimated breeding values) unbiased by genetic trend in the population. It was hoped that a comparison of the sire rankings obtained using the N.Z. Dairy Board's method with those obtained from applying a method of sire evaluation which is known to rank sires unbiased by genetic trend (MMS), would reveal whether the sire rankings obtained by the N.Z. Dairy Board are biased. However, in Chapter 5, it was noted that because the MMS estimates of the breeding values of many sires were based on comparatively few daughter records, a definitive comparison of the MMS method and the N.Z. Dairy Board method of sire evaluation was precluded.

The method of sire evaluation operated by the N. Z. Dairy Board prior to 1969 was very similar to the herdmate comparison method of evaluation (Heidhues et al., 1961) used in the United States. According to Henderson (1973), the genetic trend in the dairy cow population of New York State is sufficient to bias seriously the herdmate comparison method of sire evaluation. On this basis, it is reasonable to expect that the method in use in New Zealand prior to 1969 would also have been biased by genetic trend. However, with respect to the method of sire evaluation now in use in New Zealand (N. Z. Dairy Board, 1970), the fact that the estimated breeding value of most young sires incorporates ancestral information (for instance, the rating of their own sire) may mean that genetic trend is not a major source of bias in the method. Attention has been repeatedly drawn to the fact that in a population which is improving genetically, estimates of progeny superiority which are based on parental information are usually overestimated. Therefore, in taking account of sire and dam information in calculating the ratings of young bulls, the N. Z. Dairy Board's method of sire evaluation may not be seriously biased by genetic trend. An important exception, however, is that of young bulls whose ratings are not based on any ancestral information. In this case, not only will comparisons between their ratings and those of long-standing sires be biased by genetic trend, but so too will comparisons with other contemporary young bulls whose ratings have been based partly on ancestral information. Unfortunately, the results presented here provide no direct evidence on these points.

Until modern computing techniques were developed, a major advantage of methods of sire evaluation similar to that currently operated by the N. Z. Dairy Board over MMS methods of evaluation, has been their lower demands on computing facilities. According to Henderson (1973), however, it is possible to apply the MMS method on a routine basis using less computer time than that required to implement the herdmate comparison method of evaluation. Even so, Keown (1974) noted that a disadvantage of the MMS method in comparison to earlier techniques, is that all sires must be re-evaluated in each analysis to maximize the numbers of lactations on which the comparisons for all sires are based.

In the present study, attempts were made to apply the MMS method of evaluation to the A.I. data (3,101 records, 111 sires, and 299 herd-years), but because of computing difficulties, the scale of the exercise was reduced to the analysis of the edited-A.I. data (2,155 records, 47 sires, and 145 herd-years). The major difficulty in the case of the A.I. data was in obtaining the inverse of the matrix ($S-BH^{-1}B'$) (order = 111 x 111) using the Gauss-Jordan method of matrix inversion which is based on Gaussian elimination. In general, matrices which have been obtained using least squares principles (as in the present study) are ill-conditioned. In a study of the accuracy of some widely used matrix inversion procedures for obtaining inverses of least squares matrices, Wampler (1970) showed that procedures utilizing Gaussian elimination gave unreliable results (from the point of view of large rounding errors). The procedure finally employed obtains initially an inverse using Gaussian elimination, but then takes account of the structure of the identity matrix in proceeding to improve iteratively the accuracy of the inverse matrix. Nevertheless, rounding errors still occurred in the analysis since, despite being a consequence of the sire evaluation model, the within-group sums of sire effects (i.e., $\sum_j \hat{s}_{ij}$) were not always zero, as expected. The way to overcome this problem is to employ iterative procedures to obtain the inverse matrices (e.g., Gauss-Seidel iteration) and take advantage of the fact that the within-group sum of sire effects should be zero (Henderson, 1973).

The results in Tables 5.8 and 5.10 showed that regarding herd-year effects to be fixed or alternatively, random effects had little effect on the sire rankings produced by the MMS method of sire evaluation. This implies that, either there are minimal genetic differences among herd-years, or that in the present study, sires were evenly spread over the herd-years which differed genetically.

Even though the present work has not demonstrated that genetic trend is an important source of bias in the method of sire evaluation currently used by the N.Z. Dairy Board, it would seem that the MMS method of evaluation is the more feasible approach in ranking sires unbiased by genetic trend. Based on the experience obtained in applying the MMS method of sire evaluation used here, coupled with suggestions made by Henderson (1973), the author is of the opinion that the method could be used on a routine basis in New Zealand. In the meantime, further information is required on factors such as:

(i) The most effective method of grouping the sires to overcome the bias due to genetic trend.

(ii) Whether herd-year effects should be regarded as fixed or random effects.

(iii) Ways of obtaining accurate estimates of the variance component ratios.

(iv) The suggestions made by Henderson (1973) in minimizing the amount of computer time required to implement the method.

It is hoped that the above will be the subject of a further study.

# BIBLIOGRAPHY

Acharya, R.M., and Lush, J.L. 1968. Genetic progress through selection in a closed herd of Indian cattle. J. Dairy Sci., 51 : 1059 - 1064.

Arave, C.W., Laben, R.C., and Mead, S.W. 1964. Measurement of genetic change in twelve California dairy herds. J. Dairy Sci., 47 : 278 - 283.

Arora, K.K., and Freeman, A.E. 1970. Environmental correlation between paternal half-sibs for milk and milkfat production. J. Dairy Sci., 53 : 655 (abstr.).

Bereskin, B., and Freeman, A.E. 1965. Genetic and environmental factors in dairy sire evaluation. II. Uses and limitations of deviation records and the role of the dams. J. Dairy Sci., 48 : 352 - 355.

Bereskin, B., and Lush, J.L. 1965. Genetic and environmental factors in dairy sire evaluation. III. Influence of environmental and other extraneous correlations among the daughters. J. Dairy Sci., 48 : 356 - 360.

Bodoh, G.W., and Shook, G.E. 1972. Factors affecting selection of mates of sires in artificial service. J. Dairy Sci., 55 : 851 - 855.

Brumby, P.J. 1961. The causes of differences in production between dairy herds. Anim. Prod., 3 : 277 - 294.

Burnside, E.B., and Legates, J.E. 1967. Estimation of genetic trends in dairy cattle populations. J. Dairy Sci., 50 : 1448 - 1457.

Carter, H.W. 1969. Effect of genetic trend on first summary of artificial insemination sample sires and succeeding summaries. J. Dairy Sci., 52 : 945 (abstr.).

Chapman, A.B. 1973. Selection theory and experimental results. In Proc. Animal Breeding and Genetics Symposium in honor of Dr Jay L. Lush. Am. Soc. Anim. Sci., Am. Dairy Sci. Ass. Blacksburg, Va., July 29, 1972 : 42 - 53.

Clifford, H.J. 1972. Advances in animal production - Aspects of breeding and selection. N.Z. Agric. Sci., 6 (3) : 2 - 7.

Corley, E.L., Duckwall, J.W., and Heizer, E.E. 1963. Production performance of artificially and nonartificially sired herd-mates in Wisconsin. J. Dairy Sci., 46 : 50 - 56.

Cunningham, E.P., and Henderson, C.R. 1968. An iterative procedure for estimating fixed effects and variance components in mixed model situations. Biometrics, 24 : 13 - 25.

Dickerson, G.E. 1960. Techniques for research in quantitative animal genetics. In Techniques and Procedures in Animal Production Research. Beltsville, Md : American Society of Animal Production. pp 56 - 105.

Dickerson, G.E., and Hazel, L.N. 1944. Effectiveness of selection on progeny performance as a supplement to earlier culling in livestock. J. agric. Res., 69 : 459 - 476.

Dickinson, F.N., McDaniel, B.T., Miller, R.H., and Lytton, V.H. 1969. Estimates of sire-son relationships for yield traits in the five major dairy breeds. J. Dairy Sci., 52 : 946 (abstr.).

Edey, T.N. 1966. Herd Improvement in New Zealand - Up to 1962. Bull. 1, Massey University : 143 pp.

Evans, D.A. 1969. Research on current sire procedures. Proc. N.Z. Soc. Anim. Prod., 29 : 77 - 86.

Falconer, D.S. 1955. Patterns of response in selection experiments with mice. *Cold Spring Harb. Symp. quant. Biol.*, *20* : 178 - 196.

Goodwin, K., Dickerson, G.E., and Lamoreux, W.F. 1955. A technique for measuring genetic progress in poultry breeding experiments. *Poult. Sci.*, *34* : 1197 (abstr.).

Gustafson, G. 1970. The genetic change in milk production. A study on two Swedish breeds in 1959 to 1965. *LantbrHögsk. Annlr.*, *36* : 255 - 271.

Hargrove, G.L., and Legates, J.E. 1971. Biases in dairy sire evaluation attributable to genetic trend and female selection. *J. Dairy Sci.*, *54* : 1041 - 1051.

Harville, D.A. 1968. Statistical dependence between subclass means and the numbers of observations in the subclasses for the two-way completely random classification. *J. Am. Statistical Ass.*, *63* : 1484 - 1494.

Harville, D.A., and Henderson, C.R. 1967. Environmental and genetic trends in production and their effects on sire evaluation. *J. Dairy Sci.*, *50* : 871 - 875.

Heidhues, T., Van Vleck, L.D., and Henderson, C.R. 1961. Actual and expected accuracy of sire proofs under the New York system of sampling bulls. *Z. Tierzücht. ZüchtBiol.*, *75* : 323 - 330.

Henderson, C.R. 1949. Estimation of changes in herd environment. 5 pp. (mimeo) (also abstr. *J. Dairy Sci.*, *32* : 706).

Henderson, C.R. 1953. Estimation of variance and covariance components. *Biometrics*, *9* : 226 - 252.

Henderson, C.R. 1958. Estimates of environmental trends and biases resulting from errors in age factors and repeatability. *J. Dairy Sci.*, *41* : 747 (abstr.).

Henderson, C. R. 1973. Sire evaluation and genetic trends. In Proc. Animal Breeding and Genetics Symposium in honor of Dr Jay L. Lush. Am. Soc. Anim. Sci., Am. Dairy Sci. Ass. Blacksburg, Va., July 29, 1972 : 10 - 41.

Henderson, C. R., Carter, H.W., and Godfrey, J.T. 1954. Use of the contemporary herd average in appraising progeny tests of dairy bulls. J. Anim. Sci., 13 : 959 (abstr.).

Henderson, C. R., and Carter, H.W. 1957. Improvement of progeny tests by adjusting for herd, year, and season of freshening. J. Dairy Sci., 40 : 638 (abstr.).

Henderson, C. R., Kempthorne, O., Searle, S. R., and von Krosigk, C.M. 1959. The estimation of environmental and genetic trends in records subject to culling. Biometrics, 15 : 192 - 218.

Hickman, C.G. 1971. Response to selection of breeding stock for milk solids production. J. Dairy Sci., 54 : 191 - 198.

Hickman, C.G., and Henderson, C.R. 1955. Components of the relationship between level of production and rate of maturity in dairy cattle. J. Dairy Sci., 38 : 883 - 890.

Hickman, C.G., and Freeman, A.E. 1969. New approach to experimental designs for selection studies in dairy cattle and other species. J. Dairy Sci., 52 : 1044 - 1054.

Hill, W.G. 1972a. Estimation of genetic change. I. General theory and design of control populations. Anim. Breed. Abstr., 40 : 1 - 15.

Hill, W.G. 1972b. Estimation of genetic change. II. Experimental evaluation of control populations. Anim. Breed. Abstr., 40 : 193 - 213.

Hillers, J.K., and Freeman, A.E. 1966. Two sources of genetic error in sire proofs. J. Dairy Sci., 49 : 1245 - 1248.

Jackson, R.J. 1973. Contract mating. (Unpublished paper, N.Z. Dairy Board, 26 pp.).

Jackson, R.J. 1974. Contract mating. *Unpublished report, N.Z. Herd Imp. Council, March, 1974* : 4 pp.

Kay, R.R., and M'Candlish, A.C. 1929. Factors affecting the yield and quality of milk. I. The age of the cow. *J. agric. Sci., Camb.*, *19* : 342 - 372.

Keown, J.F. 1974. Comparison of mixed model methods of sire evaluation. *J. Dairy Sci.*, *57* : 245 - 250.

Legates, J.E. 1971. Selective improvement of milk yield in dairy cattle. *Proc. 20th U.S. National Poultry Breeders' Round Table. Kansas City, Mis., May 5-6, 1971* : 116 - 128.

Legates, J.E., and Myers, R.M. 1966. Measuring genetic change in a dairy herd using a control population. *J. Dairy Sci.*, *49* : 723 (abstr.).

Lentz, W.E., Miller, P.D., and Henderson, C.R. 1969. A new method of dairy sire evaluation. *J. Anim. Sci.*, *29* : 109 (abstr.).

Lindström, U., and Maijala, K. 1972. Improving accuracy of bull dam selection. *Acta Agric. scand.*, *22* : 189 - 199.

Lörtscher, H. 1937. Variatiosstatistiche untersuchungen an leistungserhebungen in einer British-Friesian herde. *Z. Tierzücht. ZüchtBiol.*, *39* : 257 - 262.

Lush, J.L., and Shrode, R.R. 1950. Changes in milk production with age and milking frequency. *J. Dairy Sci.*, *33* : 338 - 357.

McArthur, A.T.G. 1954. The assessment of progeny tests of dairy bulls made under farm conditions. *Proc. Br. Soc. Anim. Prod.*, pp 75 - 82.

McDaniel, B.T. 1969. New developments in genetic evaluation of dairy bulls. *Talk presented to an. meet. Natl. Assoc. Anim. Breeders, Madison, Wis., August 19, 1969* : 16 pp.

McDaniel, B.T., Plowman, R.D., and Davis, R.F. 1961. Causes and estimation of environmental changes in a dairy herd. J. Dairy Sci., 44 : 699 - 707.

McDaniel, B.T., and Corley, E.L. 1966. Environmental influences on age correction factors. J. Dairy Sci., 49 : 736 (abstr.).

McDaniel, B.T., Norman, H.D., and Dickinson, F.N. 1973. Herdmate versus contemporaries for evaluating progeny tests of dairy bulls. J. Dairy Sci., 56 : 1545 - 1558.

McDaniel, B.T., and King, G.J. 1974. Milk and fat differences for registered cows sired by artificial and natural insemination. J. Dairy Sci., 57 : 112 - 120.

Macmillan, K.L. 1973. A.B. in New Zealand dairy herds. I. Losses between conception of heifer replacements and their calving. N.Z. Jl exp. Agric., 1 : 303 - 307.

Macmillan, K.L., and Watson, J.D. 1973. A.B. in New Zealand dairy herds. II. Interactions between conception rate and submission rate on the proportion of the herd reported in calf to A.B. N.Z. Jl exp. Agric., 1 : 309 - 314.

Mao, I.L., Henderson, C.R., and Miller, P.D. 1972. Intrasire regression of daughter on herdmate performance : Nature of estimators and trend of estimates. J. Dairy Sci., 55 : 845 - 850.

Miller, P.D. 1969. Relating progeny superiority to genetic trend in cattle. J. Anim. Sci., 28 : 577 - 583.

Miller, P.D., and Henderson, C.R. 1968. Seasonal age correction factors by maximum likelihood. J. Dairy Sci., 51 : 958 (abstr.).

Miller, P.D., Lentz, W.E., and Henderson, C.R. 1970. Joint influence of month and age of calving on milk yield of Holstein cows in the North eastern United States. J. Dairy Sci., 53 : 351 - 357.

Miller, R.H. 1964. Biases in the estimation of the regression of milk production on age. J. Dairy Sci., 47 : 855 - 860.

Miller, R.H. 1970. Age, breed, trait and regional variation in the regression of daughters on herdmates. J. Dairy Sci., 53 : 1461 - 1468.

Miller, R.H., and Corley, E.L. 1965. Usefulness of information on mates of sires in artificial insemination. J. Dairy Sci., 48 : 580 - 585.

Miller, R.H., Harvey, W.R., Tabler, K.A., McDaniel, B.T., and Corley, E.L. 1966. Maximum likelihood estimates of age effects. J. Dairy Sci., 49 : 65 - 73.

Miller, R.H., McDaniel, B.T., and Dickinson, F.N. 1970. Regression of mature-equivalent production on age at calving. J. Dairy Sci., 53 : 453 - 459.

Nelson, R.H. 1943. Measuring the amount of genetic change in a herd average. J. Anim. Sci., 2 : 358 (abstr.).

N.Z. Dairy Board. 1943. 19th An. Rep., 1942-43 season : 39 pp.

N.Z. Dairy Board. 1958. 34th An. Rep., 1957-58 season : 12 pp.

N.Z. Dairy Board. 1961. 37th An. Rep., 1960-61 season : 127 pp

N.Z. Dairy Board. 1969. 45th Farm Prod. Rep., 1968-69 season : 85 pp.

N.Z. Dairy Board. 1970. 46th Farm Prod. Rep., 1969-70 season : 92 pp.

N.Z. Dairy Board. 1972. 48th Farm Prod. Rep., 1971-72 season : 98 pp.

N.Z. Dairy Board. 1973. Artificial Breeding Service, Reference Catalogue : 76 pp.

N.Z. Dairy Board. 1974. 49th Farm Prod. Rep., 1972-73 season : 102 pp.

Norman, H.D., McDaniel, B.T., and Dickinson, F.N. 1971. Variation in the genetic level of herdmates for milk. J. Dairy Sci., 54 : 782 - 783 (abstr.).

Norman, H.D., McDaniel, B.T., and Dickinson, F.N. 1972. Conflicts between heritability estimates of mature equivalent and herdmate-deviation milk and fat. J. Dairy Sci., 55 : 507 - 517.

Ødegård, A.K. 1968. Relationship between progeny tests of fathers and sons. Acta Agric. scand., 18 : 143 - 144.

Ødegård, A.K., and Robertson, A. 1968. A comparison of progeny tests for milk production on fathers and sons. Acta Agric. scand., 17 : 241 - 246.

Palmer, J.E., Wilcox, C.J., Martin, F.G., Verde, O.G., and Barrantes, R.E. 1972. Genetic trends in milk production in an experiment station Jersey herd. J. Dairy Sci., 55 : 631 - 632.

Pirchner, F., and Lush, J.L. 1959. Genetic and environmental portions of the variation among herds in butterfat production. J. Dairy Sci., 42 : 115 - 122.

Powell, R.L., Spike, P.W., and Meadows, C.E. 1973. Characteristics of first lactations. J. Dairy Sci. 56 : 812 - 816.

Rae, A.L. 1971. Sire selection. (Unpublished notes, Massey University.)

Rendel, J.M., and Robertson, A. 1950. Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. J. Genet., 50 : 1 - 8.

Rendel, J.M., Robertson, A., Asker, A.A., Kishin, S.S., and Ragab, M.T. 1957. The inheritance of milk production characteristics. J. agric. Sci., Camb., 48 : 426 - 431.

Robertson, A. 1960. The progeny testing of dairy bulls - a comparison of tests on father and son. J. agric. Sci., Camb., 54 : 100 - 104.

Robertson, A., and Rendel, J.M. 1950. The use of progeny testing with artificial insemination in dairy cattle. J. Genet., 50 : 21 - 31.

Robertson, A., and Rendel, J.M. 1954. The performance of heifers got by artificial insemination. J. agric. Sci., Camb., 44 : 184 - 192.

Robertson, A., Stewart, A., and Ashton, E.D. 1956. The progeny assessment of dairy sires for milk : The use of contemporary comparisons. Proc. Br. Soc. Anim. Prod., pp 43 - 50.

Rønningen, K. 1971. A study on selectivity of mating in animal breeding. Acta Agric. scand., 21 : 116 - 120.

Searle, S.R. 1959. Use of the selection index method to account for the number of daughters and records in progeny-testing a dairy bull. J. Dairy Sci., 42 : 1666 - 1674.

Searle, S.R. 1960. Simplified herd-level age-correction factors. J. Dairy Sci., 43 : 821 - 824.

Searle, S.R. 1961a. Part lactations. III. Progeny testing with part lactation records. J. Dairy Sci., 44 : 921 - 927.

Searle, S.R. 1961b. Estimating the heritability of butterfat production. J. agric. Sci., Camb., 57 : 289 - 294.

Searle, S.R. 1962a. Age and herd effects in New Zealand dairy cow records. J. Dairy Sci., 45 : 82 - 85.

Searle, S.R. 1962b. Bull sampling programmes in the artificial breeding of dairy cattle. Proc. N.Z. Soc. Anim. Prod., 22 : 54 - 66.

Searle, S.R. 1963. Genetic studies of dairy production early in lactation. J. Dairy Sci., 46 : 1266 - 1272.

Searle, S.R. 1964a. Review of sire-proving methods in New Zealand, Great Britain and New York State. J. Dairy Sci., 47 : 402 - 413.

Searle, S.R. 1964b. Progeny-tests of sire and son. J. Dairy Sci., 47 : 414 - 420.

Searle, S.R. 1968. Another look at Henderson's methods of estimating variance components. Biometrics, 24 : 749 - 787.

Searle, S.R., and Henderson, C.R. 1959. Establishing age-correction factors related to level of herd production. J. Dairy Sci., 42 : 824 - 835.

Shannon, A.P. 1968. Advances in semen dilution. Proc. N.Z. Soc. Anim. Prod., 28 : 23 - 31.

Skjervold, H. 1963. The optimum size of progeny groups and optimum use of young bulls in A.I. breeding. Acta Agric. scand., 13 : 131 - 140.

Smith, C. 1962. Estimation of genetic change in farm livestock using field records. Anim. Prod., 4 : 239 - 251.

Specht, L.W., and McGilliard, L.D. 1960. Rates of improvement by progeny testing in dairy herds of various sizes. J. Dairy Sci., 43 : 63 - 75.

Stichbury, J.W. 1968. Genetic implications of advances in semen dilution. Proc. N.Z. Soc. Anim. Prod., 28 : 32 - 39.

Stonaker, H.H. 1953. Estimates of genetic changes in an Indian herd of Red Sindhi dairy cattle. J. Dairy Sci., 36 : 688 - 697.

Syrstad, O. 1966. Studies on dairy herd records. III. Estimation of genetic change. Acta Agric. scand., 16 : 3 - 14.

Thompson, R. 1969. Iterative estimation of variance components for non-orthogonal data. Biometrics, 25 : 767 - 773.

Tucker, W.L., Legates, J.E., and Farthing, B.R. 1960. Genetic improvement in production attributable to sires used in artificial insemination in North Carolina. J. Dairy Sci., 43 : 982 - 987.

Van Vleck, L.D. 1962. Effect of incomplete records on sire evaluation. J. Dairy Sci., 44 : 1511 - 1515.

Van Vleck, L.D. 1963. Regression of records on herdmate averages. J. Dairy Sci., 46 : 846 - 849.

Van Vleck, L.D. 1966a. Environmental correlations in sire evaluation. J. Dairy Sci., 49 : 56 - 60.

Van Vleck, L.D. 1966b. Paternal half-sib correlations between pairs in same and different herds. J. Dairy Sci., 49 : 195 - 198.

Van Vleck, L.D. 1969. Relative selection efficiency in retrospect of selected young sires. J. Dairy Sci., 52 : 768 - 774.

Van Vleck, L.D., and Henderson, C.R. 1961a. Improvement in production of New York Holsteins due to artificial insemination. J. Dairy Sci., 44 : 1328 - 1334.

Van Vleck, L.D., and Henderson, C.R. 1961b. Measurement of genetic trend. J. Dairy Sci., 44 : 1705 - 1710.

Van Vleck, L.D., Wadell, L.H., and Henderson, C.R. 1961. Components of variance associated with milk and fat records of artificially sired Holstein heifers. J. Anim. Sci., 20 : 812 - 816.

Van Vleck, L.D., Henderson, C.R., and Carter, H.W. 1962. Evaluation of sires available through planned mating. J. Anim. Sci., 21 : 30 - 32.

Van Vleck, L.D., and Carter, H.W. 1972. Comparison of estimated daughter superiority from pedigree records with daughter evaluation. J. Dairy Sci., 55 : 214 - 217.

Verde, O.G., Wilcox, C.J., Martin, F.G., and Reaves, C.W. 1972. Genetic trends in milk production in Florida dairy herd improvement association herds. J. Dairy Sci., 55 : 1010 - 1012.

Ward, A.H. 1937. The need for sire survey work. Address to Dom. Conf. of the Dairy Bd., Wellington, March 18, 1937 : 15 pp.

Ward, A.H., and Campbell, J.T. 1938. The practical application of age conversion factors to dairy cattle production (butterfat) records. J. agric. Sci., Camb., 28 : 509 - 522.

Wampler, R.H. 1970. A report on the accuracy of some widely used least squares computer programs. J. Am. Statistical Ass., 65 : 549 - 565.

Wickham, B.W. 1972. A study of factors affecting test day records of dairy cattle. M. Agr. Sci Thesis, Massey University Library : 89 pp.

# APPENDIX 1

List of the sequential sire numbers with corresponding N.Z. Dairy Board sire codes.

| Sire number | N.Z. Dairy Board sire code | Sire number | N.Z. Dairy Board sire code |
|---|---|---|---|
| 1 | 609 | 25 | 1439 |
| 2 | 679 | 26 | 1416 |
| 3 | 412 | 27 | 1429 |
| 4 | 446 | 28 | 1459 |
| 5 | 441 | 29 | 1404 |
| 6 | 425 | 30 | 2446 |
| 7 | 659 | 31 | 1495 |
| 8 | 691 | 32 | 2449 |
| 9 | 637 | 33 | 2512 |
| 10 | 445 | 34 | 463 |
| 11 | 373 | 35 | 1415 |
| 12 | 672 | 36 | 2415 |
| 13 | 362 | 37 | 2414 |
| 14 | 781 | 38 | 2453 |
| 15 | 773 | 39 | 3495 |
| 16 | 354 | 40 | 3444 |
| 17 | 1444 | 41 | 3469 |
| 18 | 799 | 42 | 3472 |
| 19 | 790 | 43 | 3424 |
| 20 | 717 | 44 | 2487 |
| 21 | 742 | 45 | 2469 |
| 22 | 735 | 46 | 4419 |
| 23 | 1430 | 47 | 3418 |
| 24 | 1450 | | |