

SELECTING FOR LACTATION CURVE SHAPE AND MILK YIELD IN DAIRY CATTLE

By

Theodore A. Ferris

A Thesis

Submitted to Michigan State University in partial fullfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Dairy Science

1981

.

.

ABSTRACT

SELECTING FOR LACTATION CURVE SHAPE AND MILK YIELD IN DAIRY CATTLE

by

Theodore A. Ferris

Efficiency of production has not been included in selection of dairy cattle. Feeding efficiency in lactating cows is greatest in the early stage of lactation followed by a gradual decline, but health care costs also follow a similar trend. Potentially, then, it may be desirable to select for cows which either produce more in the early part of lactation to take advantage of feeding efficiency, or to select for cows having lower lactation peaks to reduce stress and health care costs. This study determines whether the shape of the lactation curve can be changed, in what way, and to what extend the change would affect 305-day milk yield.

An equation by Wood, $y_t = at^b exp(-ct)$, was used to depict the shape of the lactation curve. Estimates of curve parameters, for initial yield (<u>a</u>), the ascent (b) and the decline after peak (c) were obtained for first lactations for each of 5,927 Michigan Holsteins on DHIA in 557 herds.

The model for variance component estimation included effects for herds which were absorbed into seasons and sires. Mixed model equations with Best Linear Unbiased Predictor (BLUP) solutions were used with restricted maximum likelihood estimators to compute variance components in an iterative process. The heritabilities for these lactation curve characteristics were .06, .09 and .15 for <u>a</u>, b and c, respectively. The genetic correlations for 305-day milk yield with <u>a</u>, b, c and peak yield were -.367, .397, .004 and .911, respectively and the phenotypic correlations were .17, .071, -.107 and .849, respectively.

To examine the potential of changing the shape of the lactation curve in conjunction with selecting for 305-day milk yield, selection indexes were set up for three strategies: 1) To increase the ascent to peak production and increase peak yield. This would shift more of the lactation production to the early stage where cows have higher feed efficiency and thereby, potentially increase overall efficiency of production. 2) To delay the time of peak and to decrease the slope to peak while either ignoring or considering persistency. This effort is directed toward reducing stress and health care costs in the early stage of lactation. 3) To flatten the lactation curve by decreasing the peak yield, then at the same time increase the initial yield and persistency which would make up for some of the loss in yield due to decreasing peak yield.

Results from indexes in the first strategy suggested that selecting for both an increase in ascent and peak yield was successful and did not decrease 305-day milk greatly. Sire rankings on these indexes were very similar to their rankings on 305-day milk alone. The second strategy was slightly successful in delaying time to peak and in decreasing the ascent to peak but it decreased the genetic gain in 305-day milk to between -38 and '76 lbs

Theodore A. Ferris

per generation. This is compared to a gain of 359 lbs when 305-day milk is selected alone. For indexes of the third strategy, selection resulted in flattening the lactation curve, but doing so at a great loss in genetic gain of 305-day milk. Generation gains ranged from -282 to 6 lbs. The use of indexes in the first strategy were most desirable from the standpoint of changing the shape of the curve in the desired direction without decreasing 305-day milk appreciably. Indexes in strategies two and three could possibly be useful if more weight were applied to 305-day yield. However, the desired change in the curve shape would be much slower.

DEDICATION

The thesis is dedicated to my parents, Joseph W. and Wilda S. Ferris, remembering their guidance and advice during my boyhood.

ACKNOWLEDGEMENTS

The author wishes to thank Dr. Ivan L. Mao for his advice, encouragement and support during this graduate program. I wish to express my appreciation to the members of my graduate guidance committee, Dr. J. L. Gill, Dr. W. T. Magee, Dr. C. E. Meadows and Dr. J. A. Speicher for their support and suggestions and in particular Dr. C. R. Anderson for the many discussions on computer analyses during my research program.

My appreciation goes to the other faculty members in the Department of Dairy Science who gave me encouragement during my graduate program. A special thanks to a number of friends who were very supportive and giving of their time.

A special thanks to Cindy Curtis-Coombs for her effort and assistance in editing and typing of this dissertation and Julie Drake for her help in preparations.

And finally, thanks goes to my entire family for their support during my endeavor.

iii

TABLE OF CONTENTS

Pag	e
List of Tables	.i
List of Figures	5
I INTRODUCTION	
II LITERATURE REVIEW	ŀ
II.1 Merit	ł
II.2 Mathematical Expressions of Lactation Curves 7	,
II.2.1 Work by Wood	}
II.2.2 Comparison of equations used to fit lactation curves	
II.2.2.1 Weighted vs. unweighted log-linear form of Wood's equation	•
II.2.2.2 Linear vs. nonlinear form of Wood's equation	•
II.2.2.3 Comparisons of other equations 15)
II.2.3 Problems fitting lactation curves)
II.3 Environmental Effects on Lactation Curve Character- istics)
II.4 Genetic Parameters for Lactation Curve Characteristics . 22	•
II.5 Selection Index Method	,
II.6 Genetic Progress of Using Selection Index	
III MATERIALS AND METHODS	1
III.1 Data	1

	III.1.1 Source-defining the population	37
	III.1.2 Calculation of 305-day production from test day information	38
	III.1.3 Data screening procedure	40
	III.2 Selecting the Method to Fit Individual Lactation Curves	42
	III.3 Model	47
	III.3.1 Adjusting data for age at freshening	47
	III.3.2 Equations and assumptions of model	48
	III.3.3 Absorption of herds	60
	III.4 Variance Component Estimation	62
	III.5 Heritability, Genetic and Phenotypic Correlations .	66
	III.6 Select Indexes	69
	III.6.1 Justification and strategies	69
	III.6.2 Computation of selection index criteria	72
	III.6.3 Computing genetic change and correlated genetic response	74
	III.6.4 Computing new curves after selection	75
	III.6.5 Computing and ranking sires on indexes	76
IV	RESULTS AND DISCUSSION	78
	IV.1 Test for Normality	78
	IV.2 Marquardt's Method for Least-Squares Estimation on Nonlinear Parameters	80
	IV.3 Data	83
	IV.3.1 Variance Components	92
	IV.3.2 Heritabilities, genetic and phenotypic correlations	98

Page

	IV.4 Genetic and Correlated Genetic Change	102
	IV.4.1 Changes in lactation production	119
	IV.4.2 Changes in the shape of lactation curves	142
	IV.4.3 Summary of changes caused by selection	
	indexes · · · · · · · · · · · · · · · · · · ·	161
	IV.5 BLUP Solutions for Sires	163
	IV.5.1 Ranking sires by indexes	164
V	SUMMARY AND CONCLUSIONS	167
	List of References	173

Page

.

LIST OF TABLES

Table		Page
II.4.1	Genetic and phenotypic correlations among lactation curve parameters for first lactation by Shanks et al., 1980	24
III.1.1	Frequency distribution (percent range) of first lactation records by season and age	40
III.1.2	Frequency distribution of first lactation records by age	41
III.1.3	Frequency distribution of first lactation records by sires	41
III.4.1	Average of heritabilities reported in the literature and initial ratios used for 305-day milk and lact- ation parameters	65
III.6.1	Indexes for the three strategies and their weights .	71
IV.3.1	Amount of data after each step of screening	85
IV.3.2	Total and adjusted sums of squares	88
IV.3.3	Means, standard deviations and ranges of 305-day milk yield and lactation curve parameters before adjusting for age and age squared	90
IV.3.4	Comparison of lactation curve parameter means in the present study and other studies reported in the literature for first lactation cows	91
IV.3.5	REML estimates of error and sire variance com- ponents for 305-day milk yield and lactation curve parameters	93
IV.3.6	Genetic variances (diagonals) and covariances (off-diagonals) for 305-day milk and curve parameters	94

Table

IV.3.7	Phenotypic variances (diagonals) and covariances (off-diagonals) for 305-day milk and curve parameters . 95
IV.3.8	Standardized genetic variances (diagonals) and covariances (off-diagonals)
IV.3.9	Standardized phenotypic variances (diagonals) and covariances (off-diagonals)
IV.3.10	Heritabilities, genetic and phenotypic correlations for 305-day milk yield and lactation curve parameters
IV.3.11	Heritability values reported for lactation curve parameters and 305-day milk
IV.4.1	Change in lactation curve parameters after one generation of selection for 305-day milk yield alone
IV.4.2	Change in each lactation curve parameter when selecting alone for itself
IV.4.3	Change in 305-day milk yield when selecting for lactation curve parameters alone
IV.4.4	Change in lactation curve parameters when selecting for a, b or c alone
IV.4.5	Genetic change in 305-day milk and curve parameters for various indexes after 1 generation of selection .106
IV.4.6	Percent genetic change for 305-day milk and curve parameters in various indexes relative to change when selecting for milk alone for 1 generation 109
IV.4.7	Percent genetic change in 305-day milk and curve parameters relative to change when milk or parameter is selected alone for 1 generation
IV.4.8	Generations 1, 5 and 10 genetic values for 305-day milk, peak and time of peak for some group 1 indexes and milk selected alone
IV.4.9	Generation 1, 5 and 10 genetic values for 305-day milk, peak and time of peak for group 2 indexes 125

<u>Table</u>

IV.4.10	Generation 1, 5 and 10 genetic values for 305-day milk, peak and time of peak for group 3 indexes	127
IV.4.11	Genetic values for curve parameters a, b and c for generations 1, 5 and 10 using group 1 indexes	129
IV.4.12	Genetic values for curve parameters a, b and c for generations 1, 5 and 10 using group 2 indexes .	131
IV.4.13	Genetic values for curve parameters a, b and c for generations 1, 5 and 10 using group 3 indexes .	132
IV.4.14	Discrepancies between expected and estimated 305-day milk yield and peak yield for various indexes	134
IV.5.1	Standard deviations of BLUPs for 305-day milk yield and lactation curve parameters for 150 sires	163
IV.5.2	Rank correlations between sires ranked for milk only and other indexes	165

Page

•

LIST OF FIGURES

Figure		Page
1	A hypothetical curvilinear relationship between two traits, X and Y	141
2	Change in the shape of the lactation curve when <u>a</u> is changed from 25 to 43 in Wood's equation \ldots	143
3	Change in the shape of the lactation curve when b is changed from .24 to .33 in Wood's equation .	144
4	Change in the shape of the lactation curve when c is changed from .032 to .041 in Wood's equation	145
5	Change of the shape of the lactation curve after one generation of selection on <u>a</u> alone	147
6	Change of the shape of the lactation curve after five generations of selection on <u>a</u> alone \ldots \ldots	148
7	Change of the shape of the lactation curve after ten generations of selection on <u>a</u> alone	149
8	Change of the shape of the lactation curve after one generation of selection on b alone	151
9	Change of the shape of the lactation curve after five generations of selection on b alone	152
10	Change in the shape of the lactation curve after ten generations of selection on b alone	153
11	Change in the shape of the lactation curve after one generation of negative selection on c alone .	154
12	Change in the shape of the lactation curve after five generations of negative selection on c alone.	155
13	Change in the shape of the lactation curve after ten generations of negative selection on c alone .	156

Figure

14	Change in the shape of the lactation curve after one generation of selection on index 1 milk, 10a, 1b10c and 1 peak vield	157
15	Change in shape of the lactation curves after five generations of selection on index 1 milk, 10a, lb, -10c and 1 peak yield	158
16	Change in the shape of the lactation curve after ten generations of selection on index 1 milk, 10a, 1b, -10c and 1 peak yield	159
17	Change in shape of the lactation curve after ten generations of selection in index 1 milk and 6 peak yield	160

Page

. . .

.

I INTRODUCTION

Selection of dairy sires has primarily been based on single trait evaluation of the total lactation milk production of daughters and/or butterfat yield, and in some cases, type traits. These traits are also considered in cow selection. More recently, milk, fat and overall type have been considered in an index as an alternative or supplemental method of ranking sires.

Total merit of an individual, in the strict sense, refers to the genotype for a particular trait or group of traits weighted according to their economic value. Selection index is referred to when the phenotypes of a number of traits, usually of economic importance, are considered jointly. The index of a particular individual may be defined in as many ways as there are indexes combining a number of traits by various weights.

In the broad sense, total merit of an individual represents its overall economic value genetically plus the return over costs generated by the individual. This would include total value of milk, meat and offspring minus any costs associated with the outputs. One may consider such variables as feed costs, costs of reproduction, health care costs, and loss of production due to disease and physical characteristics. Presently, it is difficult to get information on many of these traits in order to determine genetic parameters and include them in a total merit scheme.

Efficiency of production is defined as dollars of output divided by dollars of input. Cows with greater efficiency of production would produce a larger net return. Selection on 305-day milk and fat yield is essentially selecting for gross milk income per lactation. Efficiency at which lactations are produced has been ignored, partly due to the difficulty of obtaining data on inputs. However, it is known that cows utilize feed more efficiently in early lactation. Granting, that part of this efficiency is due to catabolism of body fat. Cows producing more in the early stage of lactation may be more efficient, i.e., produce the same amount of milk for less cost.

On the other hand, health care costs are typically greater in early lactation. These costs may be related to the stress associated with high production and the cow's inability to consume enough feed to meet her requirements. Then, it may be advantageous to select individuals which do not peak as high and are more persistent.

To better define merit, then, it would be more desirable to consider the efficiency of production. This efficiency is likely to be related to the manner in which a single or several lactations of a cow are produced. That is, the shape of a cow's lactation curve is probably strongly related to the magnitude of total outputs minus total inputs or overall efficiency. Also the shape of one lactation may influence the following lactations or lifetime productivity.

Without addressing the question of what would be the optimum shape of the lactation curve, one first needs to know if the shape is heritable. Then second, how can the shape be altered.

A number of studies have dealt with fitting mathematical equations to lactation curves. Several researchers have estimated

heritabilities of parameters within the equations used. To change the shape of the lactation curve, one would select for curve characteristics (parameters) of the equation along with total yield in a selection index. An added gain would occur if measurable lactation curve characteristics are more heritable and are highly correlated genetically to 305-day milk yield. Then, they can be used in a selection index to increase genetic gain in 305-day milk, as well as change the curve shape.

The objectives of this study are:

1) Compute the genetic parameters of the lactation curve characteristics and 305-day milk yield, their heritabilities, and genetic and phenotypic correlations;

2) Devise selection index criteria for lactation productivity using the curve parameters and 305-day milk yield;

3) Compare the genetic changes in 305-day milk yield and the curve characteristics achieved by these indexes with progress when selecting for milk only;

4) Compare sire rankings by these indexes and their ranking considering 305-day milk yield alone.

II LITERATURE REVIEW

The merit of a sire or a cow can be defined many ways, depending upon the traits under consideration. For the context of this study, merit will be a function of 305-day milk yield and desired change in the shape of the lactation curve. However, the optimum shape of the curve will not be defined. Productivity will be defined as the amount of milk yield achieved in a 305-day lactation by any selection index used to change milk yield and the shape of the production or lactation curve.

The shape of the lactation curve can be defined by an appropriate mathematical expression. Change in the shape will be a function of the change, due to selection, in the constants of the mathematical equation used. These concepts will be used to determine the flexibility of the shape of the lactation curve and the influence of change in shape on 305-day milk production.

The review background covered, will then include discussions on efficiency of production, which may suggest how the shape of the curve should be altered, merit, mathematical descriptions of the shape of lactation curves, selection index and genetic progress through selection.

II.1 Merit

Everett (1975) developed equations to predict differences between sires in return over investment for milk sold and percent return on investment for heifer and milking cow sales. Pearson (1976) discussed including sire's conception rate with a sire's predicted

difference dollar value (PD\$) to estimate the profitability of an ampule of semen. This was an attempt to express the joint effects of these two traits in deviations between sires. McGilliard (1978) computed net returns for genetically superior sires when considering income of milk and fat for daughters. Semen cost per ampule was included for each sire, while a number of other variables were simulated, such as conception rate, probability of female calves and age at freshening. These simulated variables were included to map out the income function, for all lactations over a number of generations, derived from the initial ampule of semen of a sire. Everett (1975) and McGilliard (1978) were computing by various methods, a more precise value of semen for a particular sire by considering the sire's genetic merit (Predicted Difference) and semen cost. This is reflected by income over semen cost. By doing so, they suggested a number of variables that influence the profitability of a sire's daughters. They did not address the sires' genetic merit for these traits.

Bakker et al. (1980) derived a profitability index for sires which included milk, fat yield and stayability, i.e., how long daughters remain in the milking herd. This is an attempt to expand the genetic merit of sires to traits other than milk, fat and type as was Pearson's (1976) work.

Shanks et al. (1978), Hansen et al. (1979) and Shanks et al. (1981) investigated the effect of selection for milk production on reproduction, health and health care costs of daughters. These studies suggest there is a positive correlation between milk

production and health care costs. However, the higher production more than paid for the cost of health problems. Shanks (1979) further computed heritabilities and genetic correlations for some health problems. Total health costs and total health disorders had heritabilities of .03, .12, .11 and .02, .11, .05 for lactations 1, 2 and 3, respectively. Mammary cost and mammary disorders both had heritabilities of .11 for first lactation cows. Heritabilities, in general, were low for health problems. The genetic correlations between first lactation mature equivalent (ME) milk and total health costs was .07, ME milk and total health disorders -.22 and ME milk and mammary costs was -.47. The highest genetic correlations with ME milk, outside of those computed to be greater than 1, were .76 with locomotion disorders and .69 with locomotion costs. Five variables associated with reproduction had genetic correlations of greater than 1 with ME milk. However, these traits had heritabilities of less than .02.

Work by Shanks et al. (1978), Shanks (1979), Hansen et al. (1979) and Shanks et al. (1981) suggest merit of sires can be expanded to encompass other traits which reflect losses or gains in economic value of their offspring. This would better indicate the productivity of daughters in terms of total output and the net income of daughters, i.e., outputs minus inputs. Efficiency of daughters can also be determined by dividing output by units of input and then put into terms of merit.

Hooven et al. (1968) found the genetic correlation between feed efficiency and milk production for their data was .92, with heritabilities of .46 for feed efficiency and .62 for milk production. Miller and Hooven (1969) further found that feed efficiency is greatest during early lactation and decreases throughout lactation. Part of this efficiency in early lactation is attributed to catabolism of body fat. With feed efficiency the greatest in early lactation, it may be desirable to select for individuals which produce more of their milk during this period of lactation.

Hansen et al. (1979) and Shanks et al. (1981) on the other hand, found health costs were greatest during higher production in early lactation. Therefore, one may want to select cows with lower peaks to reduce stress and possibly lower health costs. The above two cases would consider merit either in terms of net income or efficiency of production within a lactation. Then, the manner in which a cow produces a lactation may influence the efficiency for the overall lactation. This would lead to the importance of the shape of the lactation curve, which reflects the distribution of milk production during a lactation.

II.2 Mathematical Expressions of Lactation Curves

The curve of a typical lactation by a dairy cow can be described as having three stages. The first stage is an incline in production after freshening, followed by the second stage, peak production, which occurs 4 to 8 weeks after calving. The third stage is a steady decline after the peak.

Numerious studies have dealt with describing the shape of lactation curves for milk production in dairy cattle, and several are reviewed in the following sections.

II.2.1 Work by Wood

-- -

Wood (1967) stated that a number of factors may influence the total yield for a single lactation, but the general shape of the curve remains substantially unaltered. He believes that the shape of the curve is economically important and suggests that cows which produce at a moderate level throughout a lactation are to be preferred to those which produce much at their peak and little thereafter. But no reasons for these arguements were given.

Wood (1967) mentioned Gaines' (1927) formula as one of the first attempts to describe lactation curves by a mathematical function. Gaines' (1927) formula was:

$$Y = ae^{-Kt}$$
(II.2.1)

where y is yield to week t; e is the base of natural the logarithm and a and K are constants. This equation was an attempt to describe the decline in production after peak. The log-linear form of the equation was fit using a hand drawn approximation of the regression line. Wood also mentioned Nelder (1966) who described an inverse polynomial:

$$y_x = x/(b_0 + b_1 x + b_2 x^2)$$
 (II.2.2)

where y_x is the yield at week x; and b_0 , b_1 and b_2 are constants. Expected maximum yield occurs when x equals the square root of (b_0/b_2) and this yield is: $(2/b_0b_2) + b_1)^{-1}$.

Wood (1967) believed that because the lactation curve initially rises to a peak following calving and then declines gradually, that the shape is essentially a gamma curve:

$$y_t = at^b exp(-ct)$$
 (II.2.3)

where exp(-ct) represents the base of the natural logarithm and can be written as e^{-ct} , and y_t is the average daily yield in week t and <u>a</u>, b, and c are constants. Wood defined two other curve characteristics, each as a function of these constants: Peak yield occurs when t = b/c

and

Peak yield is $y_{max} = a(b/c)^{b}exp(-b)$.

Wood (1972) later mentioned that \underline{a} is a constant and a general scaling factor indicating the average daily yield at the start of lactation; that b is a parameter representing the rate of increase to peak yield; and that c represents the rate of decline after peak. The parameter \underline{a} will be underlined when it appears in a sentence.

Wood (1967) also took the integral of average daily yield to estimate total yield to the t-th week:

$$y_t = a_0 \int^T t^b exp(-ct) dt$$
, (II.2.4)

which can be evaluated using tables of the incomplete gamma function. Total yield is then:

$$y = a/c^{b+1} \Gamma(b+1)$$
 (II.2.5)

where Γ is the gamma function.

Because $y_t = \underline{a}$ when $t^b \exp(-ct) = 1$, then for lactations starting at the same level, Wood (1967) has suggested that the total yield, y, becomes a function of $c^{-(b + 1)}$. He defined this function as "persistency", and referred to it as the extent to which peak yield was maintained. Wood symbolized his term of persistency as S which will be used in this text to indicate $c^{-(b + 1)}$.

To estimate the parameters, Wood (1967) used a log-linear form of equation (II.2.3) which was solved by multiple linear regression:

 $\ln y_t = \ln a + b \ln t - ct \qquad (II.2.6)$ where ln symbolizes the natural logarithm. Multiple linear regression establishes the regression line with minimum residual error or sum of the squared deviations between the data points and the regression line. For regression, the equation takes the form:

 $\ln y_t = \ln a + b \ln t - ct + e_t \qquad (II.2.7)$ where e_t refers to the descrepency between the observed and estimated yield at week t. Equation (II.2.6) is deterministic, i.e., having no error, and (II.2.7) is probablistic which accounts for error of measurement. Hereafter, equations which are deterministic will be referred to as equations and those that are probabilistic will be referred to as models.

Wood's 1969 study investigated further the characteristic of (II.2.7). Noting it compared favorably to Nelder's inverse polynomial curve. He showed that it accounted for 95.4% of the variation in monthly yield against 84.4% for the inverse polynomial. This comparison was made, however, with a small set of data.

Wood (1967) used weekly samples of 859 Friesian lactations classified by parity and month of calving. The parameters <u>a</u>, <u>b</u> and <u>c</u> of (II.2.6) are evaluated for each cow's curve. Goodness-of-fit for the natural logarithm form of the equation was determined by the amount of variation of weekly yield accounted for by the function, i.e., the square of the multiple correlation coefficient (R^2) . Nelder (1966), however, argued the inability of R^2 's to indicate the best model. Nelder pointed out, that for the known model, $y = X^2$, that using values of X = 1, ..., n, a straight line fit yielding an R^2 greater than .93 could be obtained.

Using Fisher's Z-transformation of the multiple correlation coefficient, Wood (1969) showed that an analysis of variance of Z for individual cows indicated more variation between months of calving than within months. The equation accounted for 73.8 to 91.2% of the variation in log weekly yield with 82.3% as an average for these cows.

Wood (1976) suggests it is possible to estimate b and c only by reference to the whole population when monthly weights are used, because there would be too few points to provide any precision on individual cows. However, Wood remarked that the deviations of individual cows from a general equation with population values for the parameters (a, b and c) can provide an estimate for

goodness-of-fit. It would be necessary, however, to find estimates for individual cows to determine genetic parameters for a, b and c.

II.2.2 Comparison of equations used to fit lactation curves

II.2.2.1 Weighted vs. unweighted log-linear form of Wood's equation

Comparing the log-linear form of Wood's (1967) equation using weighted and unweighted regression, Shimizu and Umrod (1976) indicated the weighted regression equation provided a better fit. Equation (II.2.7) is the unweighted form.

The weighted regression equation was:

 $\ln y_t = \ln a + b \ln t - ct + e_t/y_t \qquad (II.2.8)$ where the inverse of the observed variable, y_t , is the weight. Results from Shimizu and Umrod (1976) suggest the weighted equation provided better fit in the early stage of lactation and the unweighted equation produced better fit in late lactation. This was determined by calculating mean deviations from the computed regression line for each cow. The weighted equation had slightly fewer abnormal curves. Abnormal curves being those with either a negative b or c.

II.2.2.2 Linearvs. nonlinear form of Wood's equation

Kellogg et al. (1977) investigated the assumption Wood (1967) made in using the log-linear equation. Wood (1967), by using the logarithm transformation of (II.2.3), made the assumption that as daily milk yield increased (peaked), the variance increased. Therefore, it was assumed that a logarithm transformation was needed to achieve homogeneous variance. Kellogg et al. (1977) used a nonlinear method of Marquardt (1963) to obtain deviation estimates in the untransformed form of Wood's equation (II.2.3). With 36 cows having four lactations, Kellogg and coworkers then found, with certain considerations, the scatter of data around the estimated lactation curves appeared uniform. This supports the use of a nonlinear equation and indicates logarithm transformation may not be appropriate. They also found the variance in the first month was smaller than for some later months but otherwise no differences among variances were observed. They suggest then, this supports the assumption of homogeneous variance for months 2 to 10 using untransformed data. Kellogg and coworkers' data consisted of monthly weights except for weekly averages used in the first two months of lactation. They were able to compare variances from month to month because all cows were tested close to the same times postpartum.

Kellogg and coworkers suggested that besides random variation contributing to comparison of variances over lactation curves, two other factors are involved.

"Cows have different lactation curves so individuals following different curves will differ much more at the second than the eight month. Secondly, the actual days postpartum for the second record of monthly production can range from about 35 to 70."

The authors concluded from this that there is more diversity in stages of lactation represented in early months than in later months. They also suggested that the nonlinear form of Wood's equation (II.2.3), using intrinsic nonlinear regression, accounted for both these factors.

Cobby and LeDu (1978) also fitted data to the untransformed equation (II.2.3) and compared results to fitting by unweighted leasts-squares of the logarithm form (II.2.7). With their data, they found unweighted leasts-squares accounted for 94.2% of the variation. Plotted residuals showed a positive to negative trend from week 2 to 18, and the estimated curve overestimated the data between the 2nd and 10th week of lactation. The curve estimated by the nonlinear regression fit the data better and produced residuals that were more uniformly distributed. Cobby and LeDu (1978) indicated there was an average reduction in residual mean square of 14% when using nonlinear techniques as opposed to linear regression on the logarithm transformed equation. It is noted that the reduction was due to minimizing squared deviations from y, instead of natural logarithm of y, which is the case with linear regression on the log-transformed equation. To compare MSE's, Cobby and LeDu first untransformed the residuals of the log-linear model and then computed a new MSE. Anderson (1981) alluded to the fact that this comparison is meaningless because untransformed residuals after a log-linear fit should not produce a smaller MSE when the nonlinear fit is expected to produce the minimum MSE for the untransformed data set.

Guest (1961) pointed out that for nonlinear equations which are transformed by logarithms, the appropriate weight for weighted least-squares is proportional to the square of the dependent variable. This gives an approximation of the nonlinear model. Cobby and LeDu (1978) used such a model:

$$y_{t} = \ln a + b \ln t - ct + e_{t}/y_{t}^{2}$$
 (II.2.9)

and found the weighted log-linear equation produced a curve similar to that of the nonlinear equation.

II.2.2.3 Comparisons of other equations

Further comparisons between equations were done by Yadav et al. (1977), using 745 lactation records from 249 cows in two breeds (Hariana and Friesian-Hariana crosses). Four equations were examined:

the exponential function

$$y_{+} = A \exp(-Kt)$$
 (II.2.10)

the inverse quadratic polynomial

$$y_t = t/(b_0 + b_1 t + b_2 t^2)$$
 (II.2.11)

the gamma-type equation

$$y_{t} = At^{D} \exp(-ct)$$
 (II.2.12)

the parabolic exponential function

$$y_{t} = A \exp(bt + ct^{2})$$
 (II.2.13)

Using the R-square value as a measure of fit, they found that the inverse quadratic polynomial and the gamma-type equations gave better descriptions of the lactation curves. The transformed versions of these four functions were explored by Basant and Bhat (1978) who used weekly milk production records of 1,202 Hariana cows to compare the relative efficiencies of the functions. After transforming the observed milk (yield, y_t) to allow for linear multiple regression methods the equations become respectively:

$$\ln y_{t} = \ln A - Kt$$
(II.2.14)
$$\ln y_{t} = b_{0} + b_{1}t + b_{2}t^{2}$$
(II.2.15)

ln	^y t	3	ln	A	+	Ъ	ln	t	-	ct	(11.2.10	6)
ln	у _t	*	ln	A	+	bt	: +	ct	2		(11.2.1)	7)

Using the R-square obtained as a measure for relative efficiencies of these functions, the authors concluded that for those first lactations that were 44 weeks in length, the gamma-type equation (II.2.16) fit the best, while shorter length lactations were best fitted by the parabolic exponential function (II.2.17). For lactations two through six the average weekly yield was best fit by the inverse polynomial (II.2.15).

Schneeberger (1981) used two models to estimate lactation curves for Swiss Brown cows:

$$\ln (y_i) = \ln (a) + b \ln (t_i) - ct_i + e_i \quad (II.2.18)$$

$$\ln (y_i) = \ln (a) + b \ln (t_i - t_0) - c(t_i - t_0) + e_i \quad (II.2.19)$$

where t_0 indicates the time of initiation of lactation which occurs prior to calving. Equation (II.2.19) gave smaller mean squared errors than (II.2.18).

Schaeffer et al. (1977) compared a nonlinear technique for predicting 305-day lactation production with methods using multiplication or extension factors and regression coefficients. The authors describe their equation as a one-compartment open equation which is:

$$y_{ij} = A \exp(-\beta (1 - t_0)) [1 - \exp(-\beta (1 - t_0))] / B \exp(\epsilon_{ij})$$
(II.2.20)

where y_{ij} is the amount of milk given on the i-th day of the lactation of the j-th cow; t₀ is a lag time parameter and may indicate when a cow's udder begins to lactate prior to calving; B is the slope of the lactation curve during the increasing production stage. A is associated with peak production; β is the slope during the decline in production after the peak; ε_{ij} is a residual effect which subsequently was split into: exp(ε_{ij}) = exp(γi sin (ip)) exp(e_{ij}) where i sin(ip) is a periodic effect served in the initial analysis and correspond biologically to a seasonal effect in the curve. γ represents the amount of periodic effect in a particular set

of records, and p is 2π divided by the length of the period which could differ among lactation groups.

The authors commented that the compartemental equation allows the possibility of studying the persistency in milk production after the peak production stage using parameters already in the equation. They briefly discussed estimation using nonlinear equations. remarking that it is often more difficult than from linear equations. First, observations were converted to natural logarithms to linearize the equation. Then for each day of the lactation from the 6th to the 305th day, the average and variance of the logs for milk and fat were calculated over cows in each of the 24 lactation These averages then became the observations to estimate groups. the parameters and the reciprocal of the variances used as weighting factors for the analysis. Days for which only a few records were available would have a smaller weight than those with many records. By using this method all parameters are estimated simultaneously for an entire group of cows at one time, instead of by individual cows.

Schaeffer and coworkers mentioned one drawback of this method was that biases due to differences in management, disease and/or

persistency among cows that influences β , B, t₀, p and γ are ignored. This is because application of the nonlinear equation for extending records in progress requires the assumptions that these parameters are constant for all cows in a group. Therefore, this method only estimates the parameter which is peak yield, for each cow.

Schaeffer et al. (1977) used standard errors of prediction (SEP) for comparison of the three methods. Their results indicated that the nonlinear method was similar to the multiplicative factors in Holsteins but slightly more accurate in Jerseys.

Congleton and Everett (1980a) investigated the error and bias in using the incomplete gamma function to describe lactation curves. A total of 653 lactations, each was at least 305-days long, were fitted by linear regression after a log transformation of Wood's (1967) equation (II.2.7). The authors noted the bias and error for predicting daily milk, during the first week of lactation, were high and then declined. This was similar to reports by Schneeberger (1981). When incomplete gamma curves were fitted to montly observations of daily milk over the entire 305-day period, the authors noted the error in predicting 305-day cumulative yield (183.5 Kg) was comparable to the prediction errors for the test interval and centering date methods. This was comparable to the standard error of estimation of cumulative milk (142 Kg) by O'Connor and Lipton (1960) and Everett et al. (1968) (144.1 to 154.6) using the test interval and centering date methods.

Congleton and Everett (1980b) noted that using bias, i.e., the sum of deviation between observed and expected, and root mean squares, that extension factors did not come as close to predicting cumulative yield as the incomplete gamma technique.

II.2.3 Problems in fitting lactation curves

One problem in generating lactation curves by a log-transformation of the incomplete gamma function (II.2.7) was the initial positive bias or over estimation due to curves with a predicted inflection point prior to freshening (Cobby and LeDu, 1978). Also noted by Cobby and LeDu (1978) was the difficulty in describing the initial rise in daily milk production following calving, before much information or data points are available. Congleton and Everett (1980a) mentioned that if the initial rising portion of the curve is short or information is lacking on this portion of the curve, linear regression on the log-transformed equation (II.2.7) would give a negative estimate for b. With positive values for a and c, both the t^b and e^{-ct} components will decrease with large values of t. Therefore, the curve will have a negative slope for all days in lactation, and peak production (b/c) will be estimated to have occurred before calving. The authors noted that curves of this shape were responsible for a large amount of the bias in predicting daily milk during the first week. Congleton and Everett (1980a) noted that this may not be the case with nonlinear techniques used by Kellogg et al. (1977), where they reported deviations for actual minus predicted milk yield for the first month following

freshening. Schaeffer et al. (1977), Shanks (1979) and Schneeberger (1981) made adjustments to overcome this problem while using a linear model.

If c alone is negative, estimated peak yield will occur also before calving. Negative estimates for b and c were much more common for lactations with a first test day 30 days or more after freshening than for initial tests within 10 days of calving (Congleton and Everett, 1980a).

II.3 Environmental Effects on Lactation Curve Characteristics

Wood (1969) noted significant changes in \underline{a} , b and c due to both parity and season of calving although seasonal effects were for only one year's data. Wood (1970) reported further work using records on animals having completed four or more lactations in 10 herds, between 1952 and 1964. This totaled 1,567 lactations of 336 cows by 89 sires. The constants a, b and c were classified in a hierarchy of parity, cow, sire and herd. A method for estimating components of variance and covariance for non-orthogonal data was claimed to have been used. The method and model for the analysis of variance was not discussed. Analysis of variance indicated the constants were significantly different from parity to parity within cows. The sire effect was significant (P < .05) for the constants but not for S. He remarked that the curve constants differed between cows and progeny groups but after removal of parity effect, S was unaffected. He also found 77.4% of the variation in shape (b and c) was due to parity and season of

calving, but only 5.4% due to herd differences and 17.2% due to between cows. It is noted that this adds to 100% of the variance being accounted, which is unlikely. Madalena et al. (1979) found <u>a</u>, b and c were influenced by season of calving and year by season interaction, with parity only influencing <u>a</u>, and breed type affecting only <u>a</u> and c. They found that other two way interactions (year x parity, year x breed, season x breed, season x parity, breed x parity) were not significant.

Congleton and Everett (1980b) remarked that days open significantly affected c in lactation one and both b and c for second lactations. However, they concluded parity and season of calving influence the shape of the curve more than days open for all lactations.

Schaeffer et al. (1977) noted differences due to age and season in nearly all estimates and also that the slope for the declining production stage became steeper in later lactations. This slope was steeper for cows calving in March and August than for those calving in winter months (September-February). Congleton and Everett (1980b) presented seasonal effects on the parameters in table form. They noted <u>a</u> reached a maximum in the early summer while b and c peaked in the winter. The effect of fall calving was less than reported in England by Wood (1969). The seasonal effects of 305-day milk production found by Congleton and Everett (1980b), Wood (1969), and used in USDA-DHIA factors (Normal et al., 1974) generally agree. Keown and Van Vleck (1973) on the other hand,
found cows freshening in May through August had the highest average production and those calving in January through April had the highest peak production. Wood (1969) reported daily yield decreased during winter months and was stimulated in spring. This was independent of stage of lactation.

Congleton and Everett (1980b) calculated persistency as Wood (1967) defined it, i.e., persistency (S) is $c^{-(b+1)}$. They reported that S is larger for the third (762.5) than the second lactation (628.3) while the first lactation was most persistent (898.6). The authors noted that although the slope following peak production remained relatively constant regardless of milk yield, the persistency index (S) was higher for high producing cows and herds. Kellogg et al. (1977) noted the same relationship between slope after peak and lactation number.

Higher production usually has been associated with a more rapid decline after peak as reported by Appleman et al. (1969), Gooch (1935), Lamb and McGilliard (1960), Madden et al. (1955), and Mahadevan (1951). Madelena et al. (1979) reported cross-breds had both higher production and slower decline after peak than purebred cows.

II.4 Genetic Parameters for Lactation Curve Characteristics

Shanks (1979) computed genetic correlations and heritabilities of the parameters in the logarithm form of Wood's (1967) equation (II.2.7). Individual cow's lactations were fitted, and method 3 of Henderson (1953) was used to estimate sire components of variance

and covariance for the estimation of heritabilities and genetic correlations. He used a mixed model: $y = X\beta + Zu + E$; where y is a matrix of lactation curve estimates of all cows for the parameters of Wood's equation; X and Z were the known design matricies for the fixed and random effects, respectively; β included the fixed effects of the mean and herd-year-season; u represented the random sire classes; and E was a vector of random error.

Estimates of heritabilities, genetic and phenotypic correlations for first lactations reported by Shanks et al. (1980) are in table II.4.1. Heritabilities for peak yield and c were the highest. The correlation between c, a measure of decline after peak and S, Wood's definition of persistency, was -.68. A negative correlation would be expected since a decrease in c would represent a increase in the slope after peak.

Schneeberger (1981) used a modified version of the log-linear form of Wood's (1967) equation to fit individual lactation curves (II.2.19). In his mixed model for variance components estimation, sires and error were random factors and lactation, service period group, calving season, region, herd and interval between calving and 1st recording were fixed factors. Herds were nested within region. Harvey's (1972) method was used to compute variance components.

Schneeberger (1981) used three measures of persistency as defined in methods of Johansson and Hansson (1940), which were P2:1, P3:1, and P3:2, where Pk:1 refers to the yield in k-th

Table II.4.1	Genetic and phenotypic correlations
	among lactation curve parameters
	for first lactation ^A by Shanks et al,
	1980.

	ln a	Ъ	с	S	week of peak	peak yield
ln a	.10(.01) ^B	.02(.13)	.15(.10)	19(.20)	23(.20)	.82(.04)
Ъ	49	.06(.01)	.62(.07)	33(.24)	16(.22)	.40(.09)
с	09	.76	.14(.02)	68(.23)	98(.26)	.04(.08)
S	06	06	13	.02(.01)	.94(.06)	03(.19)
weak of						
peak	19	.02	18	.90	.02(.01)	.23(.17)
peak yi	eld .65	.21	.21	03	02	.23(.02)

 A - The diagonals are the heritabilities, genetic correlations are the above diagonals and phenotypic correlations are the below diagonals. If the absolute value of the phenotypic correlations > .04 then (p < .05).

B - Values in parenthesis are standard errors.

hundred days of lactation as a percentage of yield in the 1-th hundred days. The heritability estimates for these measures of persistency ranged from .19 to .29 with the largest for P3:1. Genetic correlations were .05 to .16 for persistency measures and 305-day yield, and -.23 to -.35 for persistency measures and 100-day yield. The c was not genetically correlated with 305-day yield but positively correlated with 100-day yield. The author reported that a genetic relationship between b and c, and 305-day yield was non-exsistent. However, the genetic correlations among 305-day yield and measures of persistency ranged from .05 (305-day fat yield with P2:1) to .16 (305-day milk yield with P2:1). There were positive genetic correlations for b and c with 100-day milk yield, .24 and .29, respectively.

Schneeberger (1981) concluded that the genetic correlations between 100-day yield and c, 100-day yield and measures of persistency, 305-day yield and c, and 305-day yield and measures of persistency suggest that breeding for high yield at the beginning of the lactation would lower persistency as he measured it, while genetic improvement of the standard (305-day) lactation would not affect it negatively. These conclusions agree with those by Gravert and Baptist (1976) who found a negative genetic correlation between initial yield and persistency measured by the slope of the lactation curve. Shanks and coworkers (1980) also found a low negative genetic relationship between initial yield and S.

Shanks (1979) adjusted the early part of the lactations by modifing Shook's (1975) factors to compute yield on day six.

Using Shook's factors the yield on day six is always going to be less than the first monthly test after day six. This was done to reduce the number of atypical lactations shapes, i.e., negative b's. Shanks reported less than 1% atypical curves. Schneeberger (1981) on the other hand used a parameter t_0 , (II.2.19), for the same purpose. Schaeffer and coworkers (1977), also used t_0 , where t_0 indicates the time of initial lactation but used it in a different equation (II.2.20). This is assumed to be some point prior to freshening, where the lactation process starts. These adjustments insured a curve which increases from day one to a peak. Therefore, a negative b is not possible, i.e., an ever decreasing curve.

Shimizu and Umrod (1976) noted 34% while Schneeberger (1981) noted 22% atypical lactation curves. Schneeberger (1981) noted that for both of his models the percentage of atypical curves decreased as lactation number increased. Atypical shapes were greater for flat curves (42%). On the other hand the smallest MSE was for first lactation and greatest for second lactations. Both percentage of atypical shapes and MSE was lower in the second model (II.2.19) which included t_0 . For Schneeberger's data, the majority of the atypical curves would probably be negative c's since esitmating t_0 should eliminate most negative b's.

Schneeberger (1981) remarked that when estimates for 305-day and 100-day yield were computed by integrating the estimated lactation curve, the heritabilities were high (.4). The heritabilities

for b was .15 for milk and .12 for fat, and for c, .20 and .18 for milk and fat, respectively. A summary table of mean values for curve parameters is included in the Results and Discussions section (Table IV.3.4).

II.5 Selection Index Method

Smith (1936) first applied selection index theory to plant breeding while Hazel (1943) applied the theory to animal selection. The principle mathematical results and many of the mathematical and statistical difficulties involving the construction and use of selection indexes are discussed by Cochran (1951). Henderson (1963) provided proofs for a number of the properties of selection index criteria and also expressed the selection index procedure with matrix notation for practical computation.

Selection index refers to selecting individuals from a population based on a criterion for the purpose of making genetic gain in a single trait or a number of traits. The phenotypic observations of the particular traits of interest are combined by computed weights (b's) which will be noted as a vector by the underscore character, ~ i.e., b. This is also to differentiate it from the parameter b of Wood's equation. All vectors will be denoted as underscored lower case letters, while upper case underscored letters will represent matrices.

The goal is to predict the total merit, or aggregate genotype of an individual using the selection index. For total merit, the aggregate genotype is

T = a'g (II.5.1)

where a is an k x 1 vector of relative economic weights, g is a ~ k x 1 vector of additive genetic values expressed as deviations from their means of the k economically important traits. Equation (II.5.2) demonstrates the form of an index. This is the estimator of total merit, the selection index:

$$I = b'p = a' \mathcal{G}' \mathcal{P}^{-1} p \qquad (II.5.2)$$

where <u>b</u>' is a m x l coefficient vector which is equated to $P^{-1}Ga$ from the equation Pb = Ga, where P^{-1} is the inverse of the m x m phenotypic variance-covariance matrix P, G is a k x k matrix of genetic variances and covariances and p is an m x l observation vector of phenotypes expressed as deviations from the mean of the estimated fixed effects (y - Xb).

The selection index equation for unrelated animals can be written as $g_i = C_i p_{i}^{-1} p_i$, where for the i-th animal, g_i is the vector of additive genetic values of the traits considered, and p_i is the vector of corresponding phenotypic deviations where $E(p_i p_{i}) = P_i$, and $E(p_i g_i) = C_i$. In the case where all animals are assumed unrelated i.e., when $E(g_i g_i) = 0$ and $E(p_i p_i) = 0$ for $i \neq i'$, then the matrix form is:

To achieve the total merit model, elements of g are linearly combined by relative economic weights (a), so for a single animal, omitting subscript i, the index is then of the form of equation (II.5.2) and total merit is equation (II.5.1). The form of the selection index for which the simultaneous equation for solving b becomes

Pb = Ga. (II.5.3) Substituting C for G in equation (II.5.2) gives the general case and can be used for several purposes, some of which Henderson (1963) listed for animal breeding. They were:

- selection for a single trait, using information on the individual and certain of its relatives;
- 2) selection for two or more traits, using the individual's records;
- selection on two or more traits, and using observations on individuals as well as on relatives; and
- selection of line-crosses, using data in addition to that on the specific cross.

It is noted, that for different cases, there are modifications necessary for the diagonal and off-diagonal elements in \underline{P} and/or \underline{C} . These are dictated by the number of records for the individual and each relative, and the number of animals in each relative group, and depending also upon any inbreeding. The adjustments made for these cases are described by Henderson (1963).

Mao (1971) showed the procedure to find the form of the predictor, g, with regard to selection indexes in the context of a general linear model:

y = Xh + Zf + e

where: y is an observation vector of N x l; Y and Z are known design matrices of orders N x p and N is q, respectively; h is a p x l vector of fixed effects; f is a q x l unknown vector which contains random effects from which the solutions are of importance and are selection criteria when referring to the animal's breeding values; and e is an N x l random sampling error vector.

For the selection index process using this model, both f and e are assumed to be multivariate normally distributed, random variables, with zero means, variance-covariance matrices B and R, respectively, and E(fe') = 0.

Mao (1971) supposes that for each animal, the model underlying the k-th record of the j-th relative of the i-th trait is:

 $y_{ijk} = h_{ij} + g_{ij} + ijk$

where h_{ij} is fixed, g_{ij} is the additive genetic value, and e_{ijk} denotes any other causes of variation which include non-additive genetic, environmental and sampling variation. Then:

 $P_{ijk} = y_{ijk} - h_{ij} = g_{ij} + e_{ijk},$

where p_{ijk} is the phenotypic deviation from the mean. Referring back to one of the assumptions usually made in selection indexes, it is then assumed that g_{ij} , $g_{i'j}$, $g_{ij'}$, $g_{i'j'}$, $e_{ijk'}$, $e_{i'jk'}$, $e_{jk'k'}$, $e_{i'j'k'}$, $e_{ijk''}$, $e_{ij'k'}$ and $e_{i'j'k'}$ (where $i \neq i$ and $j \neq j$ and $k \neq k$) follow a multivariate normal distribution with means and all covariances zero except those between g's. Also it is assumed that the variance of e_{ijk} is σ_e^2 , and $\sigma_p^2 = \sigma_g^2 + \sigma_e^2$.

The use of selection index also requires the assumption that one starts with an unselected initial population with no inbreeding. Henderson (1963) pointed out some of the unsolved problems of selection index. First, the consequences of non-normality on the efficiency of an index are not known when the index is constructed as though y and T have a multivariate normal distribution. Second, the consequences of using variance and covariance estimates, in place of the parameters, upon the effectiveness of selection and estimating genetic gain are not known. And third, it is not known how indexes should be constructed to maximize genetic progress when the assumptions of normality and/or known parameters do not hold. Mao (1971) explored the consequences of using estimates in place of parameters in selection index. He remarked that the influence of sampling error upon the efficiency of an index was large, but with more data available a more effective selection index can be constructed. Also, inclusion of a correlated trait was, in general, more effective, if the genetic correlation was high-positive, the environmental correlation was high-negative and the heritability was high.

Mao (1971) summarized the optimum properties of selection indexes:

1) The correlation between total merit and the index is maximum (Kempthorne, 1957; Henderson, 1963).

2) The expectation of the squared difference between merit and the index is minimum (Tallis, 1962; Henderson, 1963).

3) The probability of selecting one of the largest sample values of total merit by selecting the largest value of the index criteria is maximum (Williams, 1962).

4) The genetic progress in any one-round selection by the index is maximum (Henderson, 1963).

The first two properties hold true regardless of the distributional properties of the index and total merit. However, (3) and (4) require assuming a multivariate normal distribution of g and p used in (II.5.1) and (II.5.2), respectively. Therefore, the index (II.5.2) is the best for ranking individuals for total merit, regardless of unequal amounts of information. It is noted that all these properties exist only when the parameter values of the variances and covariances are known.

II.6 Genetic Progress Using Selection Index

The change in total merit, T, can be represented by ΔT such that:

$$\Delta T = b_{TT} \Delta I \qquad (II.6.1)$$

is the regression of merit on the index and here ΔT is the difference in merit between the entire population before selection, i.e., u_T , the population mean, and the mean of the selected individuals. At the same time, ΔI represents the change in index values, which is the selection differential or a measure of selection applied. The linear regression coefficient of merit on the index is b_{TI} . Equating $\Delta T = E(T|I) - u_T$, and $\Delta I = I - u_I$, then (II.6.1) can be written:

$$E(T|I) = u_{T} + b_{TT} (I - u_{T}),$$
 (II.6.2)

where u_{I} is the entire population mean for the index and E(T | I) represents the conditional mean of T, given I.

Henderson (1963), defines the expected genetic progress in one cycle of truncation selection on an index as:

$$\Delta T = \frac{\sigma_{TI}}{\sigma_{I}^{2}} \quad \frac{z}{q} \sigma_{I} \qquad (II.6.3)$$

or

$$= \mathbf{r}_{\mathrm{TI}} \frac{\mathbf{z}}{\mathbf{q}} \quad \sigma_{\mathrm{T}} = \mathbf{r}_{\mathrm{TI}} \mathbf{D} \sigma_{\mathrm{T}}. \quad (\mathrm{II.6.4})$$

The ordinate, z, is from the unit normal distribution at the point of truncation for selection; q is the fraction of the original population which is kept; and r_{TI} is the correlation coefficient between the index and merit. When the distribution of the selection criterion is normal, z/q is appropriate, else D is used when the population distribution is not known or is not normal. Henderson (1963) wrote the following equations in matrix form to be solved for b:

$$\begin{bmatrix} & & & \sigma_{y_{1}}^{2} + b_{2}\sigma_{y_{1}y_{2}} + \dots + b_{N}\sigma_{y_{1}y_{N}} = \sigma_{y_{1T}} \frac{I}{\sigma_{1'T}} \\ & b_{1}\sigma_{y_{1}y_{2}} + b_{2}\sigma_{y_{2}}^{2} + \dots + b_{N}\sigma_{y_{2}y_{N}} = \sigma_{y_{2T}} \frac{\sigma_{1}^{2}}{\sigma_{TT}} \\ & \vdots & \vdots & \vdots \\ & b_{1}\sigma_{y_{1}y_{N}} + b_{2}\sigma_{y_{2}y_{N}} + \dots + b_{N}\sigma_{y_{N}}^{2} = \sigma_{y_{NT}}^{2} \frac{\sigma_{1}^{2}}{\sigma_{TT}} \\ \end{bmatrix}$$
(II.6.5)

In (II.6.5), $\sigma_{I}^{2}/\sigma_{TI}$ does not influence the proportionality of the b's and has no effect on r_{TI} and therefore can be set to 1. In matrix notation (II.6.5) becomes:

$$Pb = t (II.6.6)$$

where P is the variance-covariance matrix of y's (phenotypes); b is the N x 1 solution vector (N = number of traits) and t is the vector of σ_{yT} 's or the covariance between genetic merit and the phenotype of a trait.

Using (II.6.6) to determine b, then the expected genetic ~ progress in one cycle of truncation selection by a set of selection criterion can be computed by (II.6.3) or (II.6.4). Needed to compute ΔT in (II.6.3) are σ_{TI} and σ_{I}^{2} which computationally are:

$$\sigma_{\text{TI}} = b_1 \sigma_{y_{1\text{T}}} + \dots + \sigma_N \sigma_{y_{N\text{T}}}$$
(II.6.7)

and

$$\sigma_{I}^{2} = b_{1}^{2}\sigma_{y_{1}}^{2} + 2b_{1}b_{2}\sigma_{y_{1}y_{2}}^{2} + \dots + b_{N}^{2}\sigma_{y_{N}}^{2}$$
(II.6.8)

For (II.6.4):

$$r_{TI}^{2} = \frac{b_{1}\sigma_{y_{1}T} + \dots + b_{1}\sigma_{y_{NT}}}{\sigma_{T}^{2}}, \qquad (II.6.9)$$

where $\sigma_{\mathbf{T}}^2$ can be completed by:

$$\sigma_{\mathrm{T}}^{2} = \sum_{i}^{\mathrm{N}} \mathbf{b}_{i} \frac{2}{\mathbf{g}_{i}} + 2\sum_{i}^{\mathrm{N}} \mathbf{b}_{i} \mathbf{b}_{j} \sigma_{\mathbf{g}_{i}}^{\mathrm{G}} \mathbf{y}_{j}^{\mathrm{G}}$$
(II.6.10)

and G represents the genotypic variance of trait y_i , and b y_i is the i-th solution from Pb = t.

If selection is based on the optimum index, then:

$$\sigma_{\text{TI}} = E(\underline{a}'\underline{g}\underline{p}\underline{b}) = \underline{a}'\underline{G}\underline{b} = \underline{p}'\underline{P}\underline{b} = \sigma_{\underline{I}}^{2}, \quad (\text{II.6.11})$$

therefore $b_{\text{TI}} = \sigma_{\text{TI}}/\sigma_{\underline{I}}^{2} = 1,$
then $\Delta T = \Delta I.$

Mao (1971), using matrix notation, describes the computing formula for ΔT , the true genetic progress when using the optimum index

is

$$\Delta T = \sqrt{b' Pb} D = \sqrt{a' GP}^{-1} Ga D \qquad (II.6.12)$$

when constructed with known parameter values and with phenotypic observations on all the traits in the total index.

When population parameters are not known, as in practice,

and the optimum index is not available, then one uses estimates of the optimum weights, \hat{b} , obtained from the equation: $\hat{Pb} = \hat{Ga}$ or $\hat{Cb} = t$. Mao (1971) notes that when truncation selection is performed utilizing such an index involving \hat{b} , i.e.,

the improvement in T will be:

$$\Delta T' = r_{TI} D\sigma_{T} = (\underline{a}' \underline{G} \underline{b}' \sqrt{\underline{b}' \underline{P} \underline{b}}) D = (\underline{a}' \underline{G} \underline{p}^{-1} \underline{G} \underline{a}' \sqrt{\underline{a}' \underline{G} \underline{p}^{-1} \underline{p} \underline{P}^{-1} \underline{G} \underline{a}}) D.$$
(II.6.13)

The selection intensity for upper truncation selection in a normal distribution would be D = z/q and for lower truncation selection D = -z/q. Therefore, ΔT is the maximum attainable progress and $-\Delta T$ is the minimum. Harris (1963) stated that a population of ΔT ' values exists with upper and lower limits of $+\Delta T$ and $-\Delta T$. He further remarked that with repeated estimations, different ΔT ' values giving a "population" of ΔT ' values will be distributed closer to the ΔT or true values. This occurs as the accuracy of estimation improves.

In the practical situations, the progress from selection is estimated by subsituting estimates for the true values in ΔT of (II.6.12) to obtain the estimated gain:

$$\hat{\Delta T} = \hat{r}_{TI} \hat{D\sigma}_{T} = \sqrt{a} \hat{GP}^{-1} \hat{Ga} D. \qquad (II.6.14)$$

One of the practical uses of selection index occurs when selection is desired on an unobservable or lowly heritable trait which has a high genetic correlation with a trait of higher heritability. By selecting for the trait with a higher heritability, progress in the trait of interest will be greater than selecting for it alone. Following this notion, when selection index is used, the genetic response of a single trait within an index frequently is of interest. Van Vleck (1979) demonstrated the genetic response of an individual trait included in an index by:

$$\hat{\Delta}G_{1} = \frac{\text{Cov}(G_{1}, I)}{\hat{\sigma}_{I}}$$
(II.6.15)

where $\hat{\sigma}_{I}$ comes from (II.6.8) and Cov (G₁, I) represents the genetic correlation between trait 1 and the index:

Cov
$$(G_1, I) = b_1 \hat{\sigma}_{G_1}^2 + b_2 \hat{\sigma}_{G_1 G_2} + \dots + b_N \hat{\sigma}_{G_1 G_N}$$
 (II.6.16)

The genetic response for a trait not included in an index can be computed by substituting Cov (G_1 , I) with Cov (G_{N+1} , I) where:

Cov
$$(G_{N+1}, I) = b_1 \hat{\sigma}_{G_1 G_{N+1}} + \dots + b_N \hat{\sigma}_{G_N G_{N+1}}$$
 (II.6.17)

and N+1 refers to the first trait not included in the index.

III MATERIALS AND METHODS

III.1 Data

III.1.1 Source - defining the population

Monthly records from the Michigan Dairy Herd Improvement (DHI) population of 168,193 cows and 2,390 herds were taken for the period between August 1978 through August 1980. Records used were monthly records on first lactations, with the first test day prior to 35 days into the lactation and the last test occuring after 280 days, with the requisite that these cows be Holstein and identified by sire. Test refers to official monthly test day recording of milk and butterfat produced on that day. Any cows with a reported abortion during this record were discarded as well as cows on unofficial test. After the editorial process for the above criteria, the total useable records were 10,107 lactations.

One must note that this population is a subpopulation of all DHI cows in Michigan (1144 of the 2,390 DHI herds) and is not necessarily a true random sample of the DHI population since those animals with sire identification may constitute a superior population. It would be logical for one to suggest this if those cows sired by superior artificial insemination (AI) sires are identified more frequently than those by poorer AI sires or unidentified home bred bulls. It is also generally noted that the DHI population itself is superior to the overall population of dairy cows.

III.1.2 Calculation of 305-day production from test day information

DHI 305-day records currently are estimated by the test day interval method using daily milk weights recorded monthly. This method takes the average of the test day weights for two consecutive months and multiplies it by the number of days between these test dates. This then is the amount of milk estimated to be produced during this interval between tests. The daily milk for all days between the calving data and the first test is estimated to be the same as produced on the first test day. Likewise, if the last test occurs prior to 305-days, the daily yield estimates between that test and 305-days is computed to be the same as the amount produced at the last test day. These estimates produce a positive bias because cows are usually increasing in production in the early stage of lactation and decreasing when they are approaching 305-days. Shook (1975) presented adjustments to the test interval method for the first, second and last tests. The adjustment for the first test accounts for the usual incline to a peak around 45 days into lactation. Because a cow is normally increasing in production prior to 30 days, less milk is actually produced than is credited by the test interval method. Therefore, a Shook factor is used to adjust this part of the cows estimated production.

For the second test, an adjustment is made if the typical peak time occurs between the first and second test date, which would cut off the top of the peak. Therefore, a Shook factor

here adds to a cow's production estimate. Finally, when the last test occurs prior to 305-day and since a cow is normally declining at this point, the test interval estimate for this period would be biased upward. Therefore, a Shook factor is used to make the adjusted estimate for the last interval.

Examples of the computations of these adjustments used on this data set (Shook factors are in parenthesis) are given below: For the first lactation record with the first test of 46 lbs at 30 days in lactation, one would have

46 lbs x 30 days x (.84), giving 1159.2 lbs where .84 is the appropriate Shook factor. Then for a second test of 50 lbs at 62 days in lactation, one would have [46 + 50]/2 x 32 days x (1.01) giving 2294.72 lbs. Then for a last test of 32 lbs at 280 days and dry at 305 days, one would have 32 x 25 days x (.96) giving 768 lbs.

If a test after 305 days was reported, the interval between 305 days and the previous test was computed by interpolation. For example:

with a yield of 31 lbs at 290 days and 25 lbs at 320 days, one would have 320 minus 290 giving 30 days and 305 minus 290 giving 15 days so that:

 $15/30 \times [31 - 25] = 3$ lbs

then 25 + 3 (28 lbs) is the estimate on day 305 then:

 $[31 + 28]/2 \times 15 \text{ days} = 442.5 \text{ lbs}$.

III.1.3 Data screening procedure

In addition to the pre-requisites for records to be included [III.1.1], more records were deleted for: 1) sires having fewer than 8 daughters, 2) herds having only one sire, and 3) herds having fewer than 3 cows. This was done simultaneously.

The restriction on the number of daughters per sire was arbitrary. Herds having only one sire were deleted because sire would be confounded with herds and would not contribute to the estimation anyways. Herds with fewer than 3 cows also would not have enough degrees of freedom to contribute to the estimation of the factors in the model. The total usable records was then reduced from 10,107 to 5,927 cows after 3 rounds of deletions.

Tables III.1.1, III.1.2, and III.1.3 show the distribution of records by seasons, ages and sires.

		age in mont	hs of freshenin	<u>8</u>
Sea	son	22-30	31-36	A11
1	Jan-Feb	6.8-12.8	5.0-7.7	9.38
2	Mar-Apr	7.4-15.0	2.5-14.5	11.15
3	May-June	9.9-15.0	7.4-17.3	11.87
4	July-Aug	19.3-28.9	29.0-34.6	26.87
5	Sept-Oct	24.7-34.3	25.6-34.2	28.95
6	Nov-Dec	9.3-15.7	8.5-19.5	11.85

Table III.1.1. Frequency distribution (percent range) of first lactation records by season and age.

100.00

Age	Freq. (%)	Age	Freq. (%)
<u><</u> 18	.24	29	7.16
19	.16	30	6.61
20	.20	31	4.64
21	.79	32	4.11
22	1.85	33	3.79
23	4.98	34	2.76
24	10.77	35	2.32
25	12.82	36	1.61
26	13.10	37	1.35
27	10.20	38	.89
28	9.64		
			100.00

Table III.1.2. Frequency distribution of first lactation records by age.

Table III.1.3. Frequency distribution of first lactation records by sires.

daughters per sire	freq. of sires(N)
1-7	18
1-10	48
11-20	50
21-50	26
51-100	12
>100	15
range (4-339)	total 151

Crosstabulation of age by season indicated a similar distribution within ages across the six seasons. Table III.1.1 indicates the seasonal distribution within the two ranges of ages are very similar. For example, cows freshening in November and December make up a similar percentage within each of the two age ranges, 9.3 to 15.7% vs. 8.5 to 19.5% for age ranges of 22 to 30 and 31 to 36 months, respectively. Table III.1.2 indicates that the majority of first lactation cows freshened between 24 and 28 months of age. By crosstabulation, it was noted that this age distribution was similar for most sires. Table III.1.3 shows the number of daughters per sire, which range from 4 to 339 with only 18 sires having fewer than 8 daughters.

III.2 Selecting the Method to Fit Individual Lactation Curves

The criteria for selecting the appropriate method and model to fit individual lactation curves should be based on their compliance with the assumptions of regression analysis. Therefore, the method, be it linear regression, weighted linear regression, or nonlinear regression, plus the model used, should produce homogeneous variance with normally distributed and independent residuals. Homogeneous error variance requires equal variance regardless of magnitude of the dependent variable, y. Therefore, there is no correlation between the magnitude of y and the amount of error in estimating daily milk production by the regression line. Normality refers to a normal distribution of the residuals. Independence of residuals refers to having no correlation in magnitude or sign between residuals (autocorrelation).

When homogeneous variance does not exist among residuals but residuals are independently and normally distributed, the parameter estimates curve characteristics \underline{a} , b, and c obtained by least-squares still are unbiased and consistent (i.e., as sample size goes to infinity the variance of the estimator goes to zero), but they are no longer minimum variance unbiased estimates

(Neter and Wasserman, 1974). Several test statistics for detecting heterogeneous variance were compared by Layard (1973) and Brown and Forsythe (1974), using Monte Carlo methods. For a population with a normal distribution, Bartlett's test had more power. Those tests which were found to be more robust than Bartlett's under certain non-normal distributions were not robust to all non-normal distributions. Layard (1973) suggests a minimum of 25 points to achieve good power to determine homogeneous error variance. This means 25 cows tested on or near the same days over the entire lactation would be needed. Kellogg et al. (1977) used 36 cows having 4 lactations and having weekly milk weights for the first two months and monthly weights thereafter, to look at variance over the entire curve after a nonlinear fit had been used. Since time of sample days after parturition were similar for the 36 cows, comparing variance between cows at the same days postpartum was possible. They suggested that the variances were equal after the first month. Intuitively then, a linear fit of the same data could not also produce equal variances and, therefore, Kellogg and coworkers concluded nonlinear fit was more appropriate. However, they included the cow by lactation interaction in the error term which may have influenced the results if the interaction exists.

On the other hand, it has been generally implicitly assumed, by those who have used Wood's (1967) equation (Wood 1967; Congleton and Everett, 1980a, b; Shanks et al. 1980) that as daily milk yield increased, so did variance. Therefore, a logarithm transformation of the data was thought necessary to achieve homogeneous variances across the entire lactation curve.

In this study, however, it was not possible to test for homogeneous error variance because:

1) Grouping cows by similar test days over the lactation, as Kellogg and coworkers did, would not be practical because days into lactation at test dates would be the same only for cows freshening at the same time in the same herd and therefore, few cows could be grouped.

2) Individual cows have only 10-12 tests, where a number of consecutive daily tests would be needed at different times postpartum to test for homogeneous variance within a lactation for a single cow.

Independence of errors refers to the assumption that there is no autocorrelation. Further, Kendall and Buckland (1971) defined autocorrelation as "correlation between members of series of observations ordered in time or space." The occurence of autocorrelation in a least-squares model may produce a number of important consequences (Neter and Wasserman, 1974): First, though the parameter estimates are unbiased, they no longer have the property of minimum variance and may be inefficient thus making the reliability of the estimates dubious. Second, the use of mean square error may seriously underestimate or overestimate the variance of the error term. Third, the least-squares procedure may greatly underestimated the true standard deviation of the estimated regression coefficient. Fourth, confidence intervals may not be valid.

The correlation between residuals for monthly test measurements of each cow after fitting a 10 to 12 month lactation is likely to trivial. There is little reason to suggest the residuals, after fitting each cow, would follow some repetitive sequence over a lactation. For this study, it is assumed that the 30 days between tests breaks up any autocorrelation between residuals. If data points were more closely related in time, then autocorrelation may be more likely to occur. In this study, it was not deemed necessary to test for autocorrelation of residuals.

The assumption that errors are normally distributed is not essential to derive point estimates of parameters but is required when making probability statements about the reliability of estimates in the form of confidence limits.

Normality of the residuals has not been tested for either the linear or nonlinear methods of fitting Wood's equation to lactations of dairy cows. If non-normality exists, tests for homogeneous variance may be in error (Brown and Forsythe, 1974).

For a lactation curve of 290 to 360 days there are 8 to 12 monthly sample points. For testing normality it is suggested by Gill (1978) and noted by Shapiro and Wilk (1965) that the W statistic developed by Shapiro and Wilk (1965) is well suited for samples of less than 50. They also noted that the W-test is sensitive to a wide range of non-normality.

Because testing for homogeneous variance was impossible for these data and autocorrelation is likely to be trivial for

monthly tests, the decision of which model to use for fitting lactation curves will be made based on results from testing for normality of residuals.

Testing for normality will be performed on two models:

 $y_{+} = \ln (a) + b \ln (t) + ct + e$

 $y_{+} = \ln (a) + b \ln (t) + ct + e/y^{2}$

where y_t is the daily yield at time t_0 . The first equation is the log-transformation of Wood's (1967) equation, and the second is the weighted form of the first using $1/y^2$ as the weight.

It is noted that the Taylor series is one method of estimating nonlinear parameters (Marquardt, 1963). For these types of equations above, the second equation is the first order approximation of the nonlinear function (Guest, 1961), which is the first degree of the Taylor series, i.e., the function plus the first derivative in the series. Therefore, without fitting the data by nonlinear regression, which would be costly, one can compare results of weighted regression, which is one step closer to nonlinear regression, to those of the simple log-linear model.

The General Linear Model (GLM) procedure of the Statistical Analysis System (SAS), (Barr et al. 1979), using weighted regression, will be used for the fitting of the two equations and testing of normality. The Shapiro-Wilk (1965) W statistic will be used on each of 500 randomly chosen cows to test the residuals for normality. Individual cows will be tabulated by probability levels (P) of having non-normally distributed residuals. A probability level of P < .25 will be used. Levels lower than P < .25 allow for larger type II error, i.e., accepting a set of residuals as normal when they are not. One would expect at a P level of .25 that 25% of the cows, in a population with a normal distribution, would fall outside the acceptable range of normally i.e., P < .25. A binomial test will be used to determine if the observed ratio of normal to non-normal is equal to the expected ratio. The model producing the highest probability will then be the one most likely to produce normally distributed residuals.

III.3 Model

III.3.1 Adjusting data for age at freshening

The 305-day milk lactation records in DHI data files are typically adjusted for age at fresehning when used for comparisons (McDaniel et al. 1967, Mao et al. 1974). These age adjusted records are called mature equivalent records. It is possible that age would also influence the lactation curve parameters within the first lactation. Records in this data file were adjusted for age of freshening, as well as for its quadratic term, by regression analysis. The GLM procedure of SAS (Barr et al., 1969) was used for the model:

 $y_{ij} = u + b_1 Age_i + b_2 Age_i^2 + e_{ij}$ (III.3.1) where y_{ij} is the j-th observation of the i-th age for any of the

dependent variables, i.e., 305-day milk yield or the lactation curve parameters, <u>a</u>, b, c, time of peak yield, peak yield or S.

The residuals from this regression procedure become the new y values of 305-day milk and the curve parameters adjusted for age at freshening.

This adjustment is valid only when there is not a significant interaction between age and the factors in the subsequent model for variance component analysis, i.e., herd, season and sire effects or when the correlations are simply to be removed and no bilogical interpretation of age and age² is desired. A crosstabulation of data indicated that the ages of daughters within sires appeared to be distributed similarly for most sires. Also, ages within seasons were distributed similarly (Table III.1.1) and it was assumed that ages within the 557 herds would be similar for most of the herds.

III.3.2 Equations and assumptions of model

For a model describing each of the variables of interest. The equation used will be:

 $y_{ijkm} = u + h_i + f_j + s_k + e_{ijkm}$

where:

 y_{ijkm} is the residual after the corresponding observation was adjusted for age of freshening and age of freshening squared, for the k-th sire in the j-th season in the i-th herd from a population of first lactation cows on Michigan DHI, having their sires identified, and lactating between July 1978 and August 1980, of either the <u>a</u>, b or c of Wood's (1967) equation $(y_t) = at^b$ exp(-ct), time of peak yield (b/c), peak yield $(a(b/c)^b exp(-b))$, S, 305-day milk yield or any of the pairwise combinations of these variables;

u is the mean of the named fixed effects;

- h_1 is the effect of the i-th herd, i = 1, ..., 557;
- f, is the effect of the j-th season in which a cow freshened; j = 1, ..., 6 which represents six seasons combining the months of January and February, March and April, May and June, July and August, September and October, and November and December;

s _k is the effect of the k-th sire; k = 1,, 150; and e is the residual effect associated with y ijkm
Factors h, and f, are assumed to be fixed, while y_{ijkm}
s and e $ijkm$ are assumed to be random. Further assumptions include:
 E(y) = Xb and the variance of y = V = Z G Z' + R; Var (e) = R = Iσ_e² (n is the number of observations) which implies that the e's associated with each observation of y are not correlated with other e's and that each e is independently drawn from the same population with mean zero, variance σ_e² and independence, i.e., no correlation between residuals;
 3) Normal distribution of residuals; 4) Cov(s, e) = 0, which implies no correlation between e and
the random factor, s;
5) $Var(s) = G = I_{150}\sigma_s^2$ which implies that there is no covariance between s's i.e., no additive genetic relationship and inde- pendent sampling between s's, and that each s is drawn from the same population with mean zero and variance σ^2 :
6) The sire effect, s, is normally distributed;
7) No correlation between the ranking of s_k and the number of
 8) Two and three-way interactions i.e., h by f, h by s, f by s and h by f by s are trivial and negligible.
Sire and season effects are of primary interest while herds
are considered a nuisance factor.
Converting to matrix form one obtains:
y = Xb + Zu + e
where:
<pre>y is the observation vector on either a, b, c, 305 day yield, peak yield, time of peak yield and S or any of the pairwise combinations of these values after adjustment for age. x is an n x p incidence matrix, where n = 4818 cows and P is the sum of 557 herds, 6 seasons and one column for u. It contains l's and 0's corresponding to the presence or absence of the observations in the herd and season classes, and for each</pre>

- observations in the herd and season classes, and for each observation a 1 in the column for u. b is a vector of length 564 containing the unknown constants of the fixed effects. b' = [u h₁ ... h₅₅₇ f₁ ... f₆].

Z is a 4,818 x 150 incidence matrix containing 1's and 0's corresponding to the presence or absence of observations within each sire. u is a 150 by 1 vector of non-observable random effects for s, $\tilde{u} = [s_1 \dots s_{150}].$ e is a 4,818 by 1 vector of non-observable random residuals corresponding to y. Noting then: $e \sim \text{NID} (0, I \sigma_e^2)$ $\tilde{E}(e) = 0$ $E(\tilde{y}) = \tilde{x}b$ $E(\tilde{g}) = \tilde{0}$ $Cov(s, \tilde{e}) = 0$ $Var(\tilde{y}) = V = Z G Z' + R, where$ $G = var(s) = E(ss^{\dagger})$ $= E[\underbrace{s}_{\sim} - E(\underbrace{s}_{\sim})] [\underbrace{s}_{\sim} - E(\underbrace{s}_{\sim})]'$ $\begin{bmatrix} s_1 \\ 1 \end{bmatrix} \begin{bmatrix} s_1 & s_2 & \dots & s_{150} \end{bmatrix}$ = | ¹ s₂ . . ^s150 $\sigma_{s_{1}, 2}$ $\sigma_{s_{2}}^{2}$. ^σs1, 150 σ⁻ ^s1 σ_s1, 2 , 150 = ^s150 150 and R = Var(e) = E(ee') = E[e - E(e)][e - E(e)]'



It is assumed for s and e all covariances are zero and that both have homogeneous variance i.e., $I\sigma_s^2$ and $I\sigma_e^2$ then:

 $G = \begin{cases} \sigma_{s}^{2} & 0 & \cdots & 0 \\ 0 & \sigma_{s}^{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_{s}^{2} \\ 0 & 0 & \cdots & \sigma_{s}^{2} \\ \end{bmatrix} = I_{150} \sigma_{s}^{2}$ 150×150

and

Then the variance-covariance matrix for all random factors can be written

$$\begin{array}{c} \operatorname{Var} \left(\begin{array}{c} y \\ s \\ e \\ e \end{array} \right) & \left(\begin{array}{c} v \\ -n \end{array} & \begin{array}{c} Z \sigma_s^2 \\ s \\ z^* \sigma_s^2 \end{array} & \begin{array}{c} I \\ 150 \sigma_s^2 \end{array} & 0 \\ I \\ -n \sigma_e^2 \end{array} & 0 \\ I \\ -n \sigma_e^2 \end{array} \right) \\ \end{array}$$

To illustrate the model, a hypothetical example of 10 cows was used:

Data:

.

Sire	Herd	Season	305-day milk
1	1	1	14266
2	1	1	15984
2	2	1	18067
1	1	2	15332
1	2	3	13367
1	3	2	16691
1	2	1	17605
3	1	2	16525
2	1	2	16001
3	3	2	15785

Then the data layout of a 10 cow example cross-classified for herds, seasons and sires is:

			S	Ĺre				Sire				
		1	2	3	no.			1	2	3	no.	
	1	2	2	1	5		1	2	2	0	4	
Herd	2	2	1	0	3	Season	2	2	1	2	5	
	3	1	0	1	2		3	1	0	0	1	
	no.	5	3	2	10		no.	5	3	2	10	

			Herd									
				1	2	3	•	no.				
Season			L 2 3	2 3 0	2 0 1	0 2 0)	4 5 1	_			
		I	10.	5	3	2		10	•			
Then	:											
y ž						X ~						
	(14,266 15,984 18,067 15,332 13,369 16,691 17,605 16,525 16,001 15,985	-		1 0 1 0 0 1 1 0	0 0 1 0 1 0 0 0	0 0 0 0 1 0 0 0 1	1 1 0 0 0 1 0 0	0 0 1 0 1 0 1 1 1	0 0 0 1 0 0 0 0 0			

b ∼	
$ \begin{pmatrix} u \\ h_1 \\ h_2 \\ h_3 \\ f_1 \\ f_2 \\ f_3 \end{pmatrix} $	

•

+

Z ~ ŭ e ~ + (s1 s2 (1 (e111) e112 e212 e121 e231 e321 e123 e123 e122 e313 $\begin{bmatrix} s_3 \end{bmatrix}$ + lo

•

The normal equations are then:

$$\begin{pmatrix} \mathbf{X}^{\mathsf{'}}\mathbf{X} & \mathbf{X}^{\mathsf{'}}\mathbf{Z} \\ \mathbf{Z}^{\mathsf{'}}\mathbf{X} & \mathbf{Z}^{\mathsf{'}}\mathbf{Z} \\ \mathbf{Z}^{\mathsf{'}}\mathbf{Z}^{\mathsf{'}}\mathbf{Z} \\ \mathbf{Z}^{\mathsf{'}}\mathbf{Z} \\ \mathbf{Z}^{\mathsf{'}}\mathbf{Z$$

10	5	3	2	4	5	1	5	3	2
5	5	0	0	2	3	0	2	2	1
3	0	3	0	2	0	1	2	1	0
2	0	0	2	0	2	0	1	0	1
4	2	2	0	4	0	0	2	2	0
5	3	0	2	0	5	0	2	1	2
1	0	1	0	0	0	1	1	0	0
5	2	2	1	2	2	1	5	0	0
3	2	1	0	2	1	0	0	3	0
2	1	0	1	0	2	0	0	0	2



The variance of y = V = Z G Z' + R:

	[1	0	0)	$\int \sigma_s^2$	0	0)	1	0	0	1	1	1	1	0	0	0]
	0	1	0	0	σ_{s}^{2}	0	0	1	1	0	0	0	0	0	1	0
	0	1	0	0	Ő	σ_{s}^{2}	0	0	0	0	0	0	0	1	0	1
	1	0	0	۲.	3 x	3	(3 x	, 10
V =	1	0	0			-										
	1	0	0													
	1	0	0													
	0	0	1													
	0	1	0													
	0	0	1													
	1	.0 x	3 ′													

•

 $= \begin{pmatrix} \sigma_{s}^{2} + \sigma_{e}^{2} & 0 & 0 & \sigma_{s}^{2} & \sigma_{s}^{2} \\ 0 & \sigma_{s}^{2} + \sigma_{e}^{2} & \sigma_{s}^{2} & 0 & 0 \\ 0 & \sigma_{s}^{2} & \sigma_{s}^{2} + \sigma_{e}^{2} & 0 & 0 \\ \sigma_{s}^{2} & 0 & 0 & \sigma_{s}^{2} + \sigma_{e}^{2} & \sigma_{s}^{2} \\ \sigma_{s}^{2} & 0 & 0 & \sigma_{s}^{2} & \sigma_{s}^{2} + \sigma_{e}^{2} \\ \sigma_{s}^{2} & 0 & 0 & \sigma_{s}^{2} & \sigma_{s}^{2} + \sigma_{e}^{2} \\ \sigma_{s}^{2} & 0 & 0 & \sigma_{s}^{2} & \sigma_{s}^{2} \\ \sigma_{s}^{2} & 0 & 0 & \sigma_{s}^{2} & \sigma_{s}^{2} \\ \sigma_{s}^{2} & 0 & 0 & \sigma_{s}^{2} & \sigma_{s}^{2} \\ \sigma_{s}^{2} & 0 & 0 & \sigma_{s}^{2} & \sigma_{s}^{2} \\ \sigma_{s}^{2} & \sigma_{s}^{2} & \sigma_{s}^{2} & \sigma_{s}^{2} \\ \sigma_{s}^{2$ σ_{s}^{2} 0
0 σ_{s}^{2} σ_{s}^{2} $\sigma_{s}^{2} + \sigma_{e}^{2}$ σ_{s}^{2} 0 σ<mark>2</mark>. 0 0 σ_{s}^{2} σ_{s}^{2} 0 0 0 $\begin{array}{cccc} 0 & 0 \\ 0 & 0 \\ \sigma_{s}^{2} & 0 \\ \sigma_{s}^{2} & 0 \\ \sigma_{s}^{2} & 0 \\ \sigma_{s}^{2} + \sigma_{e}^{2} & 0 \\ 0 & \sigma_{s}^{2} + \sigma_{e}^{2} \\ 0 & 0 \\ 0 & \sigma_{s}^{2} \\ \end{array}$ $\sigma_{s}^{2} + \sigma_{e}^{2}$ σ2 σs2 0 $\sigma_{s}^{2} + \sigma_{e}^{2}$

	Formu	lating	the m	Lxed mc	del	equ	ations	then,	x' ^{R⁻¹x}	deve	lopes
from:				x'							
	1 0 0 1 0 0	1 1 1 0 0 1 0 0 1 1 0 0 0 0	1 0 0 0 1 0	1 1 0 0 1 0 0 1 0 0 0 1 1 0	1 0 1 0 1 0	1 0 0 0 1 0	1 0 0 0 1 0	1 0 0 1 0 1 0			
	`				_R −1			7 x	: 10		
	$\int \frac{1/\sigma_e^2}{e}$	0	0	0		0	0	0	0	0	0
	0	1/o	2 0 e 0	0		0	0	0	0	0	0
	0	0	1/o	2 0		0	0	0	0	0	0
	0	0	0	$1/\sigma_e^2$	1	0	0	0	0	0	0
x	0	0	0	0	1/	′σ ² e	0	0	0	0	0
	0	0	0	0		0	$1/\sigma_{e}^{2}$	0	0	0	0
	0	0	0	0		0	0	1/0 ² e	0	0	0
	0	0	0	0		0	0	0	$1/\sigma_e^2$	0	0
	0	0	0	0		0	0	0	0	1/σ <mark>2</mark> e	0
	0	0	0	0		0	0	0	0	0	$1/\sigma^2_{ej}$
	$ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\$	1 0 1 0 0 1 1 0	0 0 1 0 1 0 1 0 0 0	x 0 0 0 0 0 1 0 0 0 1	1 1 0 0 1 0 0 0		0 0 1 0 1 1 1 1	0 0 0 0 1 0 0 0 0 0 0 0 0			

. *

	0 5 3 2 4 5 1	5 5 0 2 3 0	3 0 3 0 2 0 1	2 0 0 2 0 2 0	4 2 0 4 0	5 3 0 2 0 5 0	1 0 1 0 0 0 1	1/σ ² e	
Similarly,									
Z'R ⁻¹ X	:= (5 3 2	2 2 1	2 1 0	1 0 1	2 2 0	2 1 2	$ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} 3 \times 7 $	

then $X'R^{-1}Z$ having the dimensions of 7 x 10, 10 x 10 and 10 x 3 becomes:

.

 $\begin{pmatrix} 5 & 3 & 2 \\ 2 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 2 & 0 \\ 2 & 1 & 2 \\ 1 & 0 & 0 \end{pmatrix} \quad 1/\sigma_e^2$

and finally
$$Z'R^{-1}Z + G^{-1}$$
 is:

$$\begin{pmatrix}
5 + 1/\sigma_{s}^{2} & 0 & 0 \\
0 & 3 + 1/\sigma_{s}^{2} & 0 \\
0 & 0 & 2 + 1/\sigma_{s}^{2}
\end{pmatrix}$$
 $1/\sigma_{e}^{2}$

where \tilde{g}^{-1} is:

$$\begin{pmatrix} 1/\sigma_{s}^{2} & 0 & 0 \\ 0 & 1/\sigma_{s}^{2} & 0 \\ 0 & 0 & 1/\sigma_{s}^{2} \\ & & & s \end{pmatrix}$$
the right hand side is:

78,108 49,041 32,476 65,922 80,334 13,369 77,263
49,041 32,476 65,922 80,334 13,369 77,263
32,476 65,922 80,334 13,369 77,263
65,922 80,334 13,369 77,263
80,334 <u>13,369</u> 77,263
<u>13,369</u> 77,263
77,263
50 052
50,052
32,310

The partitioned mixed model equations are now:

1

•

$$\begin{pmatrix} X'R^{-1}X & XR^{-1}Z \\ \tilde{Z'R}^{-1}X & \tilde{Z'R}^{-1}Z + G^{-1} \\ \tilde{Z'R}^{-1}X & \tilde{Z'R}^{-1}Z + G^{-1} \\ \end{array} \begin{pmatrix} b \\ \tilde{s} \\ \tilde{s} \\ \tilde{s} \end{pmatrix} = \begin{pmatrix} X'R^{-1}y \\ \tilde{Z'R}^{-1}y \\ \tilde{s} \\ \tilde{s} \end{pmatrix}$$

and they are multiplied by R, and G^{-1} is multiplied by σ_e^2/σ_s^2 , the variance ratio for error and sire components.

	(159,625)	78,108	49,041	32,476	65,922	80,334	13,369	77,263	50,052	32,310
		+			n					
	<u> </u>	Ŀ	ч _ ч	Ъ,	, L	f,	Ŀ,	ls l	s ₂	3
	2	1-	0	1	0	2	0	0	0	$2+\sigma_e^2/\sigma_s^2$
	e	2	1	0	2	1	0	0	3+σ ² /σ ²	0
	S	2	2	1	2	2	1	5+σ ² /σ ² s	0	0
		0	Ч	0	0	0	1	н	0	0
	S	e	0	2	0	2	0	2	1	2
	4	2	2	0	4	0	0	7	7	0
	2	0	0	2	0	2	0	1	0	1
••• a)	3	0	m	0	2	0	-1	2	1	0
en have	5	S	0	0	2	e	0	5	2	1
Ve the	10	S	æ	7	4	Ś	Ч	Ś	e	5

where \mathbb{R}^{-1} was cancelled from both sides leaving a ratio of σ_e^2/σ_s^2 in the diagonal of the Z'Z portion. These equations will now yield Best Linear Unbiased Prediction solutions (Henderson, 1975) for sires only.

It is noted then that the mixed model equations are equivalent to the normal equations of Generalized Least-squares for the fixed effects. In this example the herds were not absorbed as will be done for the large data set used for the variance component analysis.

III.3.3 Abosorption of herds

To solve the mixed model equations, the nuisance factor, herds, will be absorbed into the effects for season and sire. This will be accomplished by using a FORTRAN program which absorbs one herd at a time using a row by colum technique while setting up equations pertaining to seasons and sires.

In setting up the normal equations:

(x'x	x'z)	(b)	(x'y)
Z'X	z'z)		Z'y

herds will be absorbed into six seasons and 150 sires reducing the X'X and X'Z and Z'X portions to 6 by 6, and 6 by 156 and 156 by 6 respectively, while leaving the Z'Z portion 150 by 150 for sires. The non-unique solutions of the fixed season effects are \tilde{b} and the unique estimates for sires are \tilde{u} .

The algorithm is as follows. First, data is sorted by herds, then sires are sorted within herd. Then, the following computations are done within each herd and summed across herds. For absorbing right hand side (RHS) terms in X'y of herds into those of seasons and Z'y of sires:

Absorbing herds into season RHS = season sums - (number of cows in season * herd sums/number of cows in herd).

Absorbing herds into sire RHS = sire sums - (number of daughters of sire * herd sums/number of cows in herd).

Noting that sire, season and herd sums refer to the sums of observations on a trait of a sire, sums of observations for cows within a season and sums of observations for cows within a herd.

For absorbing portions of X'X for herds into portions of X'Xfor seasons, X'Z for seasons by sires and Z'Z for sires:

Absorbing herds into seasons (X'X, diagonals) = Number of cows in season - (number of cows in season) 2 /number of cows in herd.

Absorbing herds into season (X'X, off diagonals) = -Number of cows in season i * number of $\tilde{c}ows$ in season i'/number of cows in herd, for i \neq i'.

Absorbing herds into season by sire X'Z = Number of daughters of a sire in a particular season - (number of cows in the season * number of daughters of the sire/number of cows in herd).

Absorbing herds into sire (Z'Z, diagonals) = number of daughters of sire - (number of daughters of sire) 2 /number of cows in herd.

Absorbing herds into sire by sire (Z'Z, off diagonals) = - Number of daughters of sire i * number of daughters of sire i'/number of cows in herd, for $i \neq i'$.

After absorption of one herd, the column and row for that herd is zeroed out for the next herd. The herd, sire and season sums are set to zero. With this procedure, only one pass of the data is required to complete the absorption and set up the normal equations. The resulting normal equations will have only six seasons and 150 sires, leaving a 156 by 156 coefficient matrix and a 156 by 1 vector for each trait and trait pair. Setting up the mixed model Equations, then, requires only the addition of $G^{-1}\sigma_e^2/\sigma_s^2$ to the random portion (Z'Z) for sires prior to solving for \tilde{b} and \tilde{u} .

III.4 Variance Component Estimation

An iterative restricted maximum likelihood (REML) procedure using solutions from mixed model equations (MME) will be used to compute variance components (Mao, 1981). Some desirable characteristics of REML are: 1) when MME solutions are used in maximum likelihood equations, non-negative estimates of variance components result. 2) the restricted maximum likelihood procedure maximizes the random portion of the likelihood which is invariant to the fixed effects in a mixed model. It does not assume that the fixed effects are known, as in maximum likelihood (ML), and therefore the estimates computed are unbiased. A reduction in degrees of freedom must accompany the estimation of the fixed effects. 3) REML can be used in iterative computations.

From the MME (III.4.1) b and u will be computed for each trait

$$\begin{bmatrix} \mathbf{b} \\ \mathbf{\hat{u}} \\ \mathbf{\hat{u}} \end{bmatrix} = \begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{\tilde{z}}'\mathbf{X} & \mathbf{\tilde{z}}'\mathbf{Z} \\ \mathbf{\tilde{z}}'\mathbf{X} & \mathbf{\tilde{z}}'\mathbf{Z} + \mathbf{\tilde{g}}^{-1}\ell \end{bmatrix}^{-} \begin{pmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{\tilde{z}}'\mathbf{y} \\ \mathbf{\tilde{z}}'\mathbf{y} \end{pmatrix}$$
(III.4.1)

where ℓ is a diagonal matrice of the variance ratio of $\hat{\sigma}_{e}^{2}/\hat{\sigma}_{s}^{2}$, and \tilde{G}_{e}^{-1} is the inverse of the relationship matrix for sires. Let \tilde{C}_{e} be the generalized inverse of the coefficient matrix

• ---

The REML estimators are

$$\hat{\sigma}_{e}^{2} = (\underline{y}'\underline{y} - \underline{\tilde{b}}'\underline{X}'\underline{y} - \underline{\tilde{u}}'\underline{Z}\underline{y}) / [n - r(\underline{X})] \qquad (III.4.3)$$
where \tilde{b} and \hat{u} are MME solutions.

When only one random factor is involved as in the present model

$$V(\hat{u}_s) = \mathcal{G} \sigma_s^2$$
(III.4.4)

and

$$\hat{\sigma}_{s}^{2} = [\hat{u}_{s}'\hat{u}_{s} + \hat{\sigma}_{e}^{2}(trC_{ss})]/q_{s} \qquad (III.4.5)$$

where q is the number of classes in the random sire effect. The estimators in III.4.3 and III.4.5 will be non-negative.

REML lends itself to iterative computaion because \hat{u} and \hat{b} depend on ℓ ; $\hat{\sigma}_s^2$ relies on \hat{u} , $\hat{\sigma}_e^2$ and ℓ ; $\hat{\sigma}_e^2$ relies on \hat{u} , \hat{b} , and ℓ ; and $\hat{\sigma}_e^2$ and $\hat{\sigma}_s^2$ are needed to compute new estimates of ℓ . To begin the iterative process, initial values for the variance ratios, ℓ , will be based on the heritability estimates found in the literature for the traits of interest (Table III.4.1). The first computation involves solving (III.4.1) for \hat{b} and \hat{u} , then computing $\hat{\sigma}_e^2$ in (III.4.3) and $\hat{\sigma}_s^2$ in (III.4.5). The tr \mathcal{C}_{ss} in (III.4.5) is the trace of the sire portion of the generalized inverse in (III.4.2) and q_s is 150 for the total number of sires. This process is then continued by replacing ℓ with the new ratio of $\hat{\sigma}_e^2/\hat{\sigma}_s^2$ and recomputing the REML estimators until the current and new ratio are not very different, i.e., converge.

To speed up the convergence of the iterative process, three times the difference between the current and newly computed ratio will be added to the current ratio instead of replacing the current

ratio with the new ratio. The iteration process will stop when the difference between the current and new ratio is less than .2.

For the equation (III.4.3), y'y must be adjusted not only for the mean but for herds because herds have been absorbed for other terms in (III.4.3). Therefore, y'y becomes the total sums of squares minus sums of squares for the mean minus sums of squares due to herds. The denominator is n - r(X) or the total number of cows - number of herds - number of seasons + 1, i.e.,

4818 - 557 - 6 + 1 = 4256.

The covariances between traits will be computed from the variance of the sum of each pair of traits and the variances of the two traits using the equation

Cov(i, i') =
$$\frac{1}{2}[V(i + i') - V(i) - V(i')]$$
 for $i \neq i'$.
(III.4.6)

The initial variance ratios for these new paired traits will be the average of the variance ratios of the two traits making up the paired traits. Table III.4.1 contains these initial variance ratios. An example of computing the variance ratio from the heritability of a trait is:

$$h^{2} = .25 = \frac{\hat{4(\sigma_{s}^{2})}}{\hat{\sigma}_{s}^{2} + \hat{\sigma}_{s}^{2}}$$
, set $\hat{\sigma}_{s}^{2} = 1$,

then

$$\frac{4(1)}{1+\sigma_e^2} = .25,$$

and

SO

$$\frac{4}{.25} - 1 = \hat{\sigma}_{e}^{2} = 15,$$
$$\hat{\sigma}_{e}^{2}/\hat{\sigma}_{s}^{2} = 15/1 = 15.$$

Using these initial ratios and the iterative process, more precise estimates of the variance components can be computed for the data set used than without iteration.

In order to keep the total sums of squares for traits and trait totals for sires within the significant digit computation range of the computer, scaling down of the magnitude of some traits was done. Those traits having large values were scaled down by division.

Table III.4.1 Average of heritabilities reported in the literature^A and initial ratios used for 305-day milk and lactation parameters

		Milk	a	Ъ	с	b/c	Peak	S
Trait	Heritability			Rat	los $(\hat{\sigma}_{s}^{2})$	$\hat{\sigma_{s}^{2}}$		
Milk ^B	.25	15	22	22	18	35	16	35
ln a	.10		39	39	29	66	24	66
Ъ	.10			39	30	- 66	24	66
с	.17				22	56	19	56
b/c time peak	of .02					200	39	100
Peak	.23						16	39
S	.02							200

A - Values are from Schneeberger (1981) and Shanks et al. (1980).
B - Milk is 305-day milk yield.

From (III.4.2) the variance of estimation for the fixed effects and the variance of prediction and variance of error of prediction for the random effects can be computed for the BLUPs. These estimates are:

Because F is a constant vector; $V(\hat{F} - F) = V(\hat{F})$, and the variance of error of estimation equals variance of estimation.

III.5 Heritability, Genetic and Phenotypic Correlations

Lush (1940) defined heritability in the "narrow sense" as the proportion of the total variance in a trait that is attributed to the average or additive effects of genes. He defined heritability in the "broad sense" as the fraction of total variance due to genetic variance, which contains variance due to additive effects plus variance due to dominance and epistatic effects. In the literature, heritability usually refers to that in the "narrow sense".

Heritability will be estimated for each of the parameters in Wood's (1967) equation, <u>a</u>, b, and c, plus peak yield, time of peak yield, S and 305-day milk yield by:

$$h^{2} = \frac{4\sigma_{s}^{2}}{\sigma_{s}^{2} + \sigma_{e}^{2}}$$
 (III.5.1)

The denominator $\sigma_s^2 + \sigma_e^2$ is the phenotypic variance after the variance due to named fixed effects in the model, which were adjusted for age, have been removed. With this method there are two possible sources of bias:

1) Epistatic bias (Dickerson, 1969) and

2) Ratio bias (Kendall and Stuart, 1969).

The expectation of the estimation of heritability is then, $E(h^2) = (h^2 + epistatic bias)(1 + ratio bias)$

The general formula for the approximate standard error of the ratio of variance components will be used to compute the standard errors of heritability (Dickerson, 1969):

$$\sigma(\hat{X}/\hat{Y}) = 4/\hat{Y}^2 \sqrt{\hat{Y}^2 V(\hat{X}) + \hat{X}^2 V(\hat{Y}) - 2\hat{X}\hat{Y} \text{ cov } (\hat{X},\hat{Y})} (III.5.2)$$

where in this case \hat{X} and \hat{Y} are the additive genetic and phenotype variances, respectively.

The covariance may be estimated as a simple linear function of the variance of i, i', and (i + i') from (III.4.6).

Then the genetic correlation between traits for sire becomes:

 $r_{G_{S}} = Cov s_{ii} / \sqrt{\sigma_{s_{i}}^{2} \sigma_{s_{i}}^{2}}$

where Cov s_{11} , is the estimate of the sire component of covariance between traits 1 and 1'. $\sigma_{s_1}^2$ is the estimate of the sire component of variance for trait 1; and $\sigma_{s_1}^2$ is the estimate of the sire component of variance for trait 1'.

The estimate for phenotypic correlations will be computed similarly by adding the component for error to the sire component. Standard errors of the genetic and phenotypic correlation estimates were calculated by procedures outlined by Grossman (1970). The equation for estimating the variance of the correlation coefficient is:

Est Var
$$(r_{\theta}) = \frac{f^2 r_{\theta}^2}{(nd)^2} \left\{ \left[\frac{a^2 (U_{11}U_{22} + U_{12}^2)}{u-1} + \frac{b^2 (V_{11}V_{22} + V_{12}^2)}{v-1} \right] / \hat{\theta}_{12}^2 + \left[\frac{a^2 U_{11}^2}{u-1} + \frac{b^2 V_{11}^2}{v-1} \right] / \hat{\theta}_{12}^2 + \left[\frac{a^2 U_{22}^2}{u-1} + \frac{b^2 V_{22}^2}{v-1} \right] / \hat{\theta}_{22}^2 + \left[\frac{a^2 U_{11}U_{12}}{u-1} + \frac{b^2 V_{11}V_{12}}{v-1} \right] / \hat{\theta}_{11} \hat{\theta}_{12} + 2\left[\frac{a^2 U_{12}U_{22}}{u-1} + \frac{b^2 V_{12}V_{12}}{v-1} \right] / \hat{\theta}_{12} \hat{\theta}_{22} + \left[\frac{a^2 U_{12}U_{22}}{u-1} + \frac{b^2 V_{12}V_{12}}{v-1} \right] / \hat{\theta}_{12} \hat{\theta}_{22} + \left[\frac{a^2 U_{12}^2}{u-1} + \frac{b^2 V_{12}^2}{v-1} \right] / \hat{\theta}_{12} \hat{\theta}_{22} \right]$$
(III.5.1)

where f^2 is 1 and 16 for the variances of the phenotypic and genetic correlations; respectively; \hat{r}_{θ}^2 is the square of the correlation between the two traits considered; U_{11} and U_{22} are the mean squares for sires for traits 1 and 2, respectively. U_{12}^2 is the square for the mean square for trait (1 + 2); V_{11} , V_{22} and V_{12}^2 represent the same mean squares for error; u and v are degrees of freedom for sires and error, respectively; $\hat{\theta}_{11}^2$, $\hat{\theta}_{22}^2$, $\hat{\theta}_{12}^2$ etc, represent the variances and covariances of traits 1 and 2; a^2 and b^2 are both 1 for the variance of the genetic correlation, and for the variance of the phenotypic correlation, a^2 is the square of the degrees of freedom for sires minus 1 and b^2 is the square of the degrees of freedom for error minus 1. One is subtracted from u and v to give unbiased estimates.

The standard errors for correlation coefficients can be computed by taking the square root of (III.5.1).

III.6 Select Indexes

III.6.1 Justification and strategies

Several strategies will be considered in getting up selection indexes to select the lactation curve characteristics and 305-day milk jointly.

1) The first strategy is an attempt to increase the amount of production in the early stage of lactation. This may be done by increasing the rate of ascent to the peak or increase the ascent and the peak yield without regard to persistency in later lactation. This strategy considers that cows are typically more efficient in utilization of feed during the early stage of lactation (Miller and Hooven, 1969). Realizing that part of this efficiency is due to mobilizing body fat (Miller and Hooven, 1969). Potentially then, more net income could be derived if cows increase in production earlier, and peak higher. These indexes (1:1 through 1:19) and their weights are listed in Table III.6.1.

2) The second strategy is an attempt to decrease the ascent to the peak or increase the time to peak in conjunction with increasing persistency. Decreasing the stress of high peak production may be possible in both cases. Hansen and coworkers (1979) found higher costs for health care in the early stage of lactation, during which time production and stress are the highest. If cows reach their peak at a more gradual rate, this may reduce stress and allow body reserves to be used more slowly. This strategy will determine if this change in shape is genetically possible

and what influence it would have on total milk yield. Obviously too, increasing persistency should have a positive effect on total yield. These indexes (2:1 through 2:9) and their weights are in Table III.6.1.

3) The third strategy is an attempt to increase initial yield (parameter <u>a</u>) and increase persistency while decreasing the peak, thus flattening the curve. This strategy considers decreasing the stress of peak production and possibly allowing body reserves to be used up more slowly while maintaining production in the later stage of lactation. This strategy is chosen, as is the second strategy, to decrease stress, but in this case by decreasing peak yield greatly as opposed to delaying it. Increasing persistency and increasing the initial yield as part of the strategy is an attempt to negate some of the loss in total production due to cutting off the peak. These indexes (3:1 through 3:6) and their weights are in Table III.6.1.

These strategies have been chosen to determine the potential of changing the shape of the lactation curve and 305-day yield through selection, using Wood's (1967) equation.

Indexes in Table III.6.1 with zeros for some of the weights are restricted indexes. These indexes attempt to restrict the genetic change in the traits with zero weights, while selecting for change in the other traits in the index. Kempthorne and Nordskog (1959) discuss the computations of restricted indexes.

After indexes are formulated, two methods will be pursued:

													-
Strategy 1	M:	11k	Ъ	I	?eak		Strategy	2	Milk	к Ъ	/c	Peak	s ^A
1:1		3	1		2		2.1	•	1		15	T	1
1:2		3	1		1		2.1		1		6	1	6
1:3		5	1		1		2.2		7		3	1	1
1:4		1	1		6		2:5		, 1		10	1	10
1:5		1	6		6				-		10	-	10
1:6		1	1	1	15								
									Milk	c c	Ъ	/c	Peak
	M111	k a		Ь	С.	Peak	2.5				10	1 5	
							2:5		1		10	10	10
1:7	1	o ^B		3	0	6	2.0		1	-	- 6	20	10
1:8	1	ŏ		6	Ō	6	2.7		T		-0	0	T
1:9	5	Ō		1	Ō	1							
	-	-		_		-			M	ii lk		с	
		Milk		Pea	1k 				-				-
							2:8			1		-1	
1:10		1		6	ò		2:9			1		-10	0
1:11		3]	_								
1:12		6]	-								
		Milk		ť)	· .	Strategy	3	Milk	: a	Ъ	Peal	k S ^A
					-		3:1		1	10	1	1	10
1:13		1		6	;		3:2		1	10	-5	-5	10
1:14		1		15	5	•	3:3		1	10	-10	-10	10
1:15		3		1	-								
1:16		6		1	-								
								Mi	llk	a	Ъ	с	Peak
		Milk		a.	Ъ	с .							
1:17		1	(0	6	0	3:4	1	_	10	1	-10	1
1:18		3	(0	1	0	3:5	1	_	10	-5	-10	-5
1:19		6	(0	1	0	3:6	1	-	10	-10	-10	-10
					· · ·					•			

Table III.6.1. Indexes for the three strategies and their weights.

A - S is c^{-(b+1)}, adjusted for a.
B - Indexes containing weights of zero are restricted indexes, where traits with zero weights are those being restricted.

1) Genetic change will be determined for each of the traits in the indexes. The correlated responses of the curve parameters (<u>a</u>, b, and c) when not included in a particular index will also be computed.

2) Best Linear Unbiased Prediction (BLUP) solutions (u) for sires from MME will be linearily combined by the weights to give a Total Merit (Index) for each sire.

Henderson (1963) noted the BLUP of $k'\beta + m'u$ is $k'\beta + m'GZ'V^{-1}$ ($y - X\beta$) and u is equal to $GZ'V^{-1}(y - X\beta)$. From MME, u is computed and therefore solving $GZ'V^{-1}(y - X\beta)$ to get u is not necessary. Also β is equal to the solutions for the fixed effects in MME, i.e., \tilde{b} . Mao (1981) notes that T = a'g = m'u the aggregate merit. Thus total merit can be computed by the linear combination of weights (a or m) and BLUP solutions (u).

III.6.2 Computation of selection index criteria

The genetic and phenotypic variance-covariance matrices used in the selection index equation

 $\mathbf{Pb} = \mathbf{Ga} \tag{III.6.1}$

will be standardized. This is done by dividing both sides of (III.6.1) by the phenotypic variance such that the diagonals (variances) in \underline{P} and \underline{G} are divided by the phenotypic variance for each trait and the off-diagonals (covariances) are divided by the product of the phenotypic standard deviations for the two traits making up the covariance. This treats both sides of (III.6.1) the same.

So, P becomes:

$$\begin{bmatrix} 1 & \hat{r}_{P_{12}} & \hat{r}_{P_{13}} \\ \hat{r}_{P_{21}} & \hat{r}_{P_{23}} \\ \hat{r}_{P_{31}} & \hat{r}_{P_{32}} & 1 \end{bmatrix}$$

where $\hat{\sigma}_{P_1}^2 / \hat{\sigma}_{P_1}^2 = 1$ and $\hat{\sigma}_{P_1P_2} / \hat{\sigma}_{P_1} \hat{\sigma}_{P_2} = \hat{r}_{P_{12}}$

which is the phenotypic correlation.

G becomes:



With the standardized \underline{P} and \underline{G} matricies, the solutions for \underline{b} in (III.6.1) will be computed. The values of \underline{b} are standardized partial regression coefficients and will be denoted as d. To compute the partial regression coefficients, d is divided by the phenotypic standard deviation for its related trait. These standard values for \underline{G} and \underline{P} will only be used to compute the d's.

It is pointed out that P and G are positive semi-definite matrices, such that all principal leasing minors have determinants greater than or equal to zero. This is necessarily the case when all heritabilities are between zero and 1 and all correlations are between -1 and +1.

III.6.3 Computing genetic change and correlated genetic response

Equation (II.6.3) expresses the genetic change of total merit as a result of the use of an index, but in this study this is not of interest. Instead, the genetic change of individual traits either within the index or not included in the index are of primary interest. In particular, 305-day milk yield and the curve parameters, <u>a</u>, b, and c are of interest, and in some cases peak yield, time of peak and S will be of interest.

Computation of genetic gain of an individual trait included in the index is as follows:

$$\hat{\Delta G}_{i} = \frac{\text{Cov}(G_{i}, I)}{\hat{\sigma}_{I}} \times z/q, \qquad (III.6.2)$$

where $Cov(G_i, I)$ represents the genetic correlation between trait i and the index, and $\hat{\sigma}_I$ comes from:

$$\sigma_{I}^{2} = b_{1}^{2} \sigma_{y_{1}}^{2} + 2b_{y_{1}y_{2}}^{2} + \dots + b_{N}^{2} \sigma_{y_{N}}^{2}, \qquad (III.6.3)$$

where $\sigma_{y_1}^2$, σ_{y_1} , etc., come from the phenotypic variance-covariance matrix. The b's in (III.6.3) are partial regression coefficients and the phenotypic variances and covariances are not standardized.

The $Cov(G_i, I)$ is computed using the genetic variance of the trait in question and its covariances with all traits in the index. If i = 1 then:

$$Cov(G_{1}, I) = E\{G_{1}[b_{1}p_{1} + b_{2}p_{2} + \dots + b_{N}p_{N}]\}$$
$$= b_{1}\hat{\sigma}_{G_{1}}^{2} + b_{2}\hat{\sigma}_{G_{1}G_{2}}^{2} + \dots + b_{N}\hat{\sigma}_{G_{1}G_{N}}^{2}.$$
(III.6.4)

For a trait not included in the index, only the covariances between the trait and all other traits in the index are used:

$$Cov(G_{N+1}, I) = b_1 \hat{\sigma}_{G_1G_{N+1}} + \dots + b_N \hat{\sigma}_{G_NG_{N+1}}$$
. (III.6.5)

Correlated responses of an unselected trait, i', when a single trait, i, is selected is computed by:

$$\Delta G_{i} = r_{G_{i}G_{i}}, \quad (\hat{\sigma}_{G_{i}}, \hat{\sigma}_{G_{i}}) \hat{\Delta u}_{i}, \quad (III.6.6)$$

where Δu_i is the change in trait i due to single trait selection of i, i.e.,

$$\hat{\Delta u}_{i} = h^{2} \hat{\sigma}_{P_{i}} z/q. \qquad (III.6.7)$$

Equations III.6.2, and III.6.4 through III.6.7 are discussed in various forms by Van Vleck (1979).

For the purpose of comparing genetic gain, z/q will remain constant and for simplicity avalue of 1 is chosen. If 5% of the sires and 90% of the cows are selected as parents, then for q and 5%, z/q = 2.1 and for q of 90%, z/q = .2, then (.2 + 2.1)/2 = 1.15. Therefore, 1 is a reasonable choice.

III.6.4 Computing new curves after selection

For each index in the three strategies, estimates for genetic change in the curve parameters \underline{a} , b and c in Wood's (1967) equation (II.2.3) will be computed for 1, 5 and 10 generations of selection.

Using these new estimates and the appropriate form of Wood's equation, be it log-linear or nonlinear, new lactation curves will be plotted for each index. The shapes of the curves generated by the indexes will be compared to the shape generated if only milk is selected.

The integrals for the new curves will be computed for 305-days into lactation and compared to the expected change in 305-day milk yield determined by $\hat{\Delta G}_{milk}$. For those indexes relating to changing the peak yield, there will be a comparison between the expected change in peak ($\hat{\Delta G}_{peak}$), and that estimated from the new values for the curve parameters in the equation:

$$Peak = a(b/c)^{b}exp(-b). \qquad (III.6.10)$$

Computing the change over 10 generations, as described above, assumes that the genetic response of individual traits as well as their correlations are linear. This may not be the case.

III.6.5 Computing and ranking sires on indexes

Computing total merit for sires for each index is done by standardizing the BLUP (\hat{u}) solutions and combining them linearily with the appropriate weights. Standardization of the BLUPs is done by dividing the BLUP for each sire by the standard deviation of the BLUPs for that trait. Standardizing puts all traits in terms of their standard deviations so that traits of low numerical value are not over-shadowed in the total merit score by traits which have high numerical values. Then, total merit is computed by:

$$T = a'g' = m'u'$$
 (III.6.11)

where a or m is a vector of weights and g^* or \hat{u}^* is a vector of standardized BLUPs. In selection index theory, \hat{u} is replacing $C' \nabla^{-1}(y-X\beta)$ in: $T = m'C' \nabla^{-1}(y-X\beta) = m'\hat{u}$.

The weights are the same as the weights listed for the indexes in Table III.6.1. These weights are relative only to changing the shape of the lactation curve and are therefore arbitrary, depending upon the direction and amount of change desired.

The ranking of the sires by their total merit for each index will be compared to their ranking on 305-day milk yield. Spearman's ranked correlation analysis will be used (Gill, 1978) to determine if selection by various indexes have significantly changed the ranking of the sires from their ranking on milk alone.

IV RESULTS AND DISCUSSION

IV.1 Test for Normality

Monthly milk weights for first lactation records for 481 randomly chosen cows were fitted to two equations below, and the residuals were tested for normality using the Shapiro-Wilk W-test (1965). The two equations are:

1) $y_{+} = \ln a + b \ln t + ct + e$

2) $y_t = \ln a + b \ln t + ct + e/y_+^2$

where y_t is daily milk yield at time t, <u>a</u>, b and c are constants, e is simple error and e/y_t^2 is the weighted regression form of error.

Each cow had 8 to 11 monthly milk weights and received a probability (P) level indicating the probability of the 8 to 11 residuals being normally distributed when the hypothesis of normality is rejected. For example, P < .25 means the probability of type I error is less than .25. A binomial test was used to determine the probability that the observed number (N) of cows with P levels less than .25 was not different than the expected number. These results for model 1 and 2 are reported in Table IV.1.1 for the random sample of 481 cows.

Table IV.1.1 Binomial probability of observed number of cows being not different from expected number having probability levels below .25.

Model		P < .25			
	Binomial Probability	Observed ^A N	۶ ⁸	Expected ^C N	
1	.0080	98.0	20.3	120	
2	.36	123	25.5	120	

A - Observed number of cows below P < .25.

B - Percent of cows below P < .25 (total = 481).

C - Expected number of cows below P < .25.

For a sample of cows from a population with normally distributed residuals, the test for normality should produce a percentage of cows having a probability level less than the chosen level of P, which is equal to the chosen P level. Therefore, if one tabulates all cows with P levels less then .25, this should include 25% of the cows.

The binomial test compares the expected N with the observed N and yields the probability that they do not differ. At P < .25one expects N = 120. In model (1), N is 98, and for model (2), N is 123. The binomial probabilities are .0080 and .36 for model (1) and (2), respectively. From the results, the model which provided the highest probability of having normally distributed residuals was model (2), an approximation of a nonlinear model.

The real concern was the comparison between the log-linear and a nonlinear model. Because model (2), the weighted linear regression model, is an approximation of a nonlinear model, and is less expensive to compute, it was used in the test for normality of residuals in place of the nonlinear model. The results suggest that the nonlinear model would be more appropriate than the log-linear model, from the standpoint of normality.

Based upon these results and findings by Kellogg et al. (1977), Cobby and LeDu (1978) and Shimizu and Umrod (1976), the nonlinear model was used. Marquardt's (1963) technique of nonlinear regression was used to fit 5,927 lactations to the nonlinear form of Wood's (1967) equation.

IV.2 Marquardt's Method for Least-Squares

Estimation on Nonlinear Parameters

Marquardt (1963) developed a maximum neighborhood method. This method utilizes the Taylor series and gradient (steepest-descent) methods of nonlinear estimation. Marquardt mentions that these two methods, when used separately, have difficulties in estimating nonlinear parameters. The maximum neighborhood method is stated to perform an optimum interpolation between the Taylor series and gradient method. The interpolation is based upon the maximum neighborhood in which the truncated Taylor series gives an adequate representation of the nonlinear model.

Marquardt states the problem as follows. Choosing a model:

 $E(y) = f(X_1, X_2, ..., X_m; \beta_1, \beta_2, ..., \beta_k)$ (IV.2.1) to be fitted to a set of data where $X_1, X_2, ..., X_m$ are independent variables and $\beta_1, \beta_2, ..., \beta_k$ are the population parameter values or <u>a</u>, b and c of Wood's (1967) equation for an individual cow or group (population) of cows. E(y) is the expected value of the dependent variable y. Data points are denoted by:

 $(y_{i}, y_{1i}, X_{2i}, ..., X_{mi})$ i = 1, 2, ..., n. (IV.2.2) The nonlinear form of Wood's (1967) equation is:

$$y_{t} = at^{D} exp(-ct) + e,$$
 (IV.2.3)

where X_1 becomes t and X_{1i} is any time, t_i , in lactation and y_i equals y_t which is daily milk yield at some time, t_i . Thus, t is the only independent value.

It is then desired to compute those estimates of the parameters which will minimize:

$$\Phi = \sum_{i=1}^{n} [\mathbf{Y}_{i} = \hat{\mathbf{Y}}_{i}]^{2}$$

where Y_i is the value of y predicted by (IV.2.1) at the i-th data point. When the function for the expected value of Y_i is linear in the β 's, the contours of constant, Φ , are ellipsoids but for the nonlinear case, they are distorted, depending upon the degree of nonlinearity. But with nonlinear models, the contours are nearly elliptical in the immediate vicinity of the minumum of Φ . Marquardt also mentions that the contour surface of Φ is very narrow in some directions and elongated in others such that the minimum lies at the bottom of a long curving trough.

Using Marquardt's notations, the equations used for iteration to a point which the residual sum of squares (ϕ) is minimized are as follows:

$$A^* = a^*_{jj}, = (\frac{a_{jj'}}{\sqrt{a_{jj}}\sqrt{a_{j'j'}}}),$$
 (IV.2.4)

where A* is a scaling matrix to scale the b-space in units of the standard deviations of the derivatives $\partial f_i / \partial b_j$ taken over the sample points i = 1, 2, ..., n. This makes the A matrix one of simple correlation coefficients of the $\partial f_i / \partial b_j$'s. a_{jj} , a_{jj} , and $a_{j'j'}$ represent various sums of squares and sums of cross-products.

The algorithm used is:

$$(A^{*} + \lambda^{r}I) \delta^{*} = g^{*}r,$$
 (IV.2.5)

representing the equation at the r-th iteration, where

scaled vector
$$g^* = A^*\delta^*_+$$
 (IV.2.6)

and

$$g^* = (g_j^*) = (\frac{g_i}{\sqrt{a_{jj}}})$$
 (IV.2.7)

and δ_t^* is the Taylor series correction

$$\delta_{j} = \frac{\delta_{j}^{*}}{\sqrt{a_{jj}}} . \qquad (IV.2.8)$$

Equation (IV.2.5) is solved for δ^{*r} and (IV.2.8) is used to obtain δ^{r} . A new trial vector:

$$b^{(r+1)} = b^{r} + \delta^{r}$$
 (IV.2.9)

will then produce a new residual sums of squares, $\phi^{(r+1)}$.

Marquardt noted it is essential to select $\lambda^{\mathbf{r}}$ such that $\phi^{(\mathbf{r+1})} < \phi^{\mathbf{r}}$ (IV.2.10)

meaning the new residual sums of squares are less then the current. A form of trial and error is used to find a value λ^r which will satisfy (IV.2.10) and produce rapid convergence of the algorithm to the least-squares values.

Marquardt's strategy was: Let v be greater than 1 (usually use 10) and let $\lambda^{(r-1)}$ denote the value of λ from the previous iteration, but initially λ^0 is equal to 10^{-2} .

Compute $\Phi(\lambda^{(r-1)})$ and $\Phi(\lambda^{(r-1)}/v)$. 1) if $\Phi(\lambda^{(r-1)}/v) \leq \Phi^{r}$; let $\lambda^{r} = \lambda^{(r-1)}/v$. 2) if $\Phi(\lambda^{(r-1)}/v) > \Phi^{r}$, and $\Phi(\lambda^{(r-1)}) \leq \Phi^{r}$; let $\lambda^{r} = \lambda^{(r-1)}$. 3) if $\Phi(\lambda^{(r-1)}/v) > \Phi^{r}$, and $\Phi(\lambda^{(r-1)}) > \Phi^{r}$; increase λ by successive multiplication by v until for some smallest w,

$$\Phi(\lambda^{(r-1)}v^{w}) \leq \Phi^{r}$$
. Let $\lambda^{r} = \lambda^{(r-1)}v^{w}$.

By this algorithm, Marquardt suggests a feasible neighborhood is obtained. The iteration converges when

$$\frac{\left|\delta_{j}^{r}\right|}{\tau + \left|b_{j}^{r}\right|} < e, \text{ for all } j,$$

for suitably small e > 0, i.e. 10^{-5} and a suitable τ , i.e. 10^{-3} . For v, a value of 10 has been found to be a good choice.

In the determination of the parameters <u>a</u>, b and c for Wood's (1967) nonlinear equation (IV.2.3), a grid search was performed for each cow to arrive at an initial best guess for the values of <u>a</u>, b and c. Then Marquardt's method was used to refine the estimates of <u>a</u>, b and c by further minimizing the sums of squares of the residuals. The whole procedure was computed using SAS NLIN procedure (Barr et al., 1979).

The partial derivatives of (IV.2.3) needed for Marquardt's procedure were:

 $\partial f_y / \partial a = t^b \exp(-ct)$ $\partial f_y / \partial b = \ln t(at^b \exp(-ct))$ $\partial f_y / \partial c = (at^b \exp(-ct)) - t.$

IV.3 Data

Table IV.3.1 describes the transition of records used at various steps of the analysis. The initial selection criterion were:

1) 1st lactation records;

2) having sire identification;

3) having 8 to 12 monthly tests;

4) lactation not coded with an abortion; and

5) one test prior 35 days and one after 280 days.

This yielded 10,107 records over the two year period, July 1978 to August 1980.

Records were further dropped for sires having fewer than eight daughters and herds with fewer than three cows. This left 5,927 records. Upon fitting these 5,927 records with the nonlinear form of Wood's (1967) equation using nonlinear regression, 887 lactations (15%) yielded negative values for parameter c. There were 14 additional lactations with negative values for b. These lactation curves were estimated to have peaked before calving and therefore would have a continuous decline from freshening. They are, therefore, classified atypical lactations.

Shimizu and Umrod (1976) reported 34 and 29% atypical lactation curves for an unweighted and weighted regression model of the logarithm form of Wood's (1967) equation. Schneeberger (1981) compared two models (II.2.18) and (II.2.19), which produced 26.6 and 25.9 atypical shapes, respectively. Schneeberger also noted the later lactation animals produced fewer atypical lactations (19-22%). If tests prior to peak are missing, then the curve would take on an ever decreasing shape and have a negative b. First lactation animals do not peak as high as later lactation animals, and they may peak earlier. Therefore, it would be likely that first

Step	Records	Herds	Sires
After selecting 1st lactations with 8 to 12 tests and sire identification.	10,107	1,114	717
After deleting sires <7 daughters and herds <3 cows (used for non- linear regression).	5,927	678	152
Fit models to test normality.	481	(random records)	
after deleting records with negative values for c.	5,040	678	151
After deleting records with nega- tive values for b.	5,026	678	151
After last deletion of herds with one sire or less than 3 cows.	4,818	557	150

Table IV.3.1. Amount of data after each step of screening.

lactation animals would have more atypical curves i.e., negative b's. Likewise, it is more likely that a first lactation animals would be increasing in production near 305-days and therefore have a negative c.

In the current study the number of atypical curves was minimized by deleting cows which did not have a test prior to 35 days into lactation. Also, first and second monthly tests were adjusted using Shook's (1975) factors, thereby accounting for the typical increase to peak. These procedures have been responsible for having fewer (15%) atypical curves in the present study than reported by Shimizu and Umrod (1976), and Schneeberger (1981). Shanks (1979) also accounted for the typical increase from parturition by using Shook's (1975) factors to compute milk yield on day six. He reported less than 1% atypical curves for all parities. Almost all of the 15% atypical curves in the present study were due to negative c values, meaning the last part of the curve was increasing. If Shook (1975) factors are used to compute the last test, then a decline is forced and a negative c is less likely to occur. In the present study this was done only if the last test occurred prior to 305 days.

The atypical records were dropped from the data leaving 5,026 records. Before computing the variance components, a total of 121 herds each having only one sire or fewer than three cows were dropped. In these herds, sires would be confounded with herd, and in herds with only two cows, one degree of freedom would be lost for herd, leaving only one degree of freedom for estimation of sire and error. This left 4,818 records, 557 herds and 150 sires.

Equations for 557 herds were absorbed in setting up mixed model equations for seasons and sires. For the computations of variance components, sequencial sums of squares were computed after removing sums of squares due to the mean and herds. These are in Table IV.3.2. These reduced sums of squares were used to compute the REML estimate for error variance $(\hat{\sigma}_e^2)$ in (III.4.3), i.e., $\hat{\sigma}_e^2 = (y'y - \hat{b}'X'y - \hat{u}X'y)/[n - r(X)],$

where $y \ y$ is the total sums of squares after removing the mean and herd effects, i.e., values in the 3rd column of Table IV.3.2. Noting that the sums of squares due to age were previously removed.

Table IV.3.3 shows the means, standard deviations and ranges for 305-day milk yield and the lactation curve parameters. These are the values before records were adjusted for age and age squared as mentioned in the method section. The average 305-day actual milk yield for the 5,927 records in this study was 14,801 lbs, which is for first lactation animals. The Michigan DHIA lactation average is currently 15,463 lbs and the Holstein breed mature equivalent (ME) is 15,416 lbs. Using the average 26 month age adjustment factors the first lactation 305-day records (ME) would be 18,192 lbs which is considerably greater then 15,463 lbs. One would expect the ME average of two yr olds to be greater than the population average if genetic progress exists. A portion of the difference may, also, be due to selecting a sub-population in which sires are identified, and requiring sires to have eight or more daughters.

	*	•	
Variables or . parameters	SST ^A	ssh ^b	SST-SSH ^C
305-day milk	36,753,873,961	14,409,519,760	22,344,354,201
a	886,152	176,247	704,405
Ъ	99.4666	18.6552	80.8107
c	.0112946	.00237596	.00891864
b/c time of peak	27,350,364	3,653,198	23,697,165
pe ak yie ld	530,856	225,713	305,143
S	13,970,035,907	2,587,541,657	11,383,494,249

.

Table IV.3.2 Total and adjusted sums of squares.

A - Sequencial sums of squares corrected for the mean.

B - Sequencial sums of squares for herds.

C - Sequencial sums of squares after mean and herds.

 $D - c^{-(b+1)}$, adjusted for a.

The mean value for <u>a</u> was 31.6. This value compares closely to those reported in the literature and summarized in Table IV.3.4, except for Schneeberger (1981). Noting here, that values for <u>a</u> are untransformed values of ln a, except for the present study, which computed <u>a</u> using the nonlinear form of Wood's equation. Schneeberger's data represents a lower producing population. The mean value for b, .212, can only be compared to the value in Schneeberger's study (.409) were, as in the present study, time was computed in days. Values for b and c are not comparable between models which are computed using time in days as opposed to time in weeks. Values for c would necessarily be smaller when time is computed in days.

Peak yield in the present study was higher than for other studies. This would be expected after comparing the 305-day production levels with those available from the other studies. The 305-day production was considerably higher than Wood (1967), 14,801 vs. 7,898 to 11,669 lbs and Schneeberger (1981) 7,132 lbs, and higher than Shanks (1979) when comparing mature equivalents, 18,192 vs. 16,465 lbs.

Time to peak was greatest in those studies with higher peaks. This is expected if Wood's (1967) equation is used. Shanks (1979) reported a late peak time of 12 weeks, but when calculated from b and c values reported, b/c was equal to 10.1 weeks.

Table IV.3.3	Means, standard deviations and ranges of 305-day milk yield and lactation
	curve parameters before adjusting for age and age squared. ^A

Variables or Parameters	mean	standard deviation	range
305-day milk yield	14800	2800	4,630-27,600
а	31.6	13.7	.264-86.0
Ъ	.212	.145	.000119-1.330
c	.00302	.00154	000479016
b/c time of peak	69.1	133	-7340-3880
pe ak yie ld	58.0	10.6	20.4-118
s ^B		3880	

A - 5,927 records were used before deletion of the atypical curves.

B - S is $c^{-(b+1)}$ adjusted for a, therefore, the mean is of residual values which is zero.

Table IV.	3.4 Comparison present st erature fc	l of lactation udy and other r first lacts	i curve para studies re ition cows.	neter means ported in th	in the e lit-	
Variables or parameters	present study ^A	Wood (1967)	Wood (1969)	Wood (1970)	Shanks .(1980)	Schneeberger (1981) ^A
đ	31.6	30.6 ^E	33.9 ^E	29.5 ^E	37.1 ^E	13.1 ^E
р	.212	.189	.150	.280	.239	.409
IJ	.00302	.0306	.030	.0360	.0237	.00497
peak (1bs)	58.0	30.0-43.0	37.2 ^D	39.7 ^D	52.8	29.3
b/c week of peak	10.0	6.17 ^D	5.00 ^D	7.77 ^D	12.0 ^B	6.96
305-day yield (lbs)	14,800	8,000-11,67	0		16,500 ^C	7,130
S		56.0-75.0			142	
A - Time was compu	uted in days.					

4 -4 -۲ ι . ċ C 111 Tabl

D - Computed from parameters given.

C - Mature equivalent. B - b/c = 10.1 weeks.

E - Values are untransformed values of ln (a)

.

IV.3.1 Variance Components

The variance components estimates for sire and error from mixed model equations and the appropriate restricted maximum likelihood estimators (REML), are shown on Table IV.3.5. All estimates have positive values. The starting values of the variance ratios for iteration in REML are in Table III.4.1. The iterations required for convergence ranged from 2 to 10. Rounds of iterations required were greatly reduced by adding three times the difference between the current ratio, and the new variance ratio, to the current ratio, instead of replacing the current ratio with the new ratio.

The genetic variances and covariances for 305-day milk yield and the lactation parameters <u>a</u>, b and c, plus time of peak, peak yield and S are in Table IV.3.6. Their phenotypic variances and covariances are in Table IV.3.7. The genetic covariance between 305-day milk and the curve parameters are all positive except for that with <u>a</u> and b/c (Table IV.3.6). The phenotypic covariances between 305-day milk and the curve parameters are all positive except for c (Table IV.3.7).

Persistency as Wood (1967) defined it is $c^{-(b+1)}$, S, but this assumed that <u>a</u> was constant for all cows. Since this is not the case, a new value for persistency was computed by adjusting for <u>a</u>, using regression in the GLM procedure of SAS (Barr et al., 1979).

Due to the drastic difference in the magnitude of the genetic and phenotypic variances and covariance for the traits, the variance-covariance matrice (P and G) were standardized. This was

done to minimize rounding error when $\underset{\sim}{P}$ was inverted for the selection index equation:

$$b = P^{-1}Ga$$
.

The standardized values for the genetic and phenotypic values are in Tables IV.3.8 and IV.3.9, respectively. Computations of these values are demonstrated in [III.6.2]. The standardized genetic variances are the heritabilities for the traits, while the standardized phenotypic variances are equal to 1. All standardized values are between +1 and -1.

Table IV.3.5.	REML estimates of error and sire variance	
	components for 305-day milk yield and	
	lactation curve parameters.	

Variables or parameters	Error variance	Sire variance
305-day milk	5,042,870	201,252
a	161.398	2.61620
b	.0179950	4.30302 E-4
c	.0191470	7.34878 E-8
b/c time of peak	5,484.95	99.84450
peak	66.9376	2.52968
S	2,641,490	25,559

A - S is, $c^{-(b+1)}$, adjusted for a.
	Milk	c	م	J	b/c	Peak	ß
Milk ^A	805,000	-1,068	14.74	.001897	-1,558	2,600	155.6
ŋ		10.46	1216	001065	20.82	-2.5512	506.8
Ą			.001721	1.00001612	2404	.09066	3.290
υ				2.941 E-7	005277	3.449 E-5	03190
b/c ^B					399.4	-43.42	1520
Peak						10.12	207.5
с S							102,200

Genetic variances (diagonals) and covariances (off-diagonals) for 305-day milk and curve parameters. Table IV.3.6

A - 305-day milk.

B - time of peak yield.

C - S is $c^{-(b+1)}$, adjusted for a.

covariances (off-diagonals) for	
Phenotypic variances (diagonals) and	305-day milk and curve parameters.
Table IV.3.7	

•

	Milk	cj	Ą	U	b/c	Peak	S
Milk ^A	5,244,000	5,016	22.07	3455	13,520	16,200	20,570
3		164.0	-1.512	01091	-258.4	13.02	-251.0
Ą			.01842	.0001520	1.942	.1991	25.72
J				1.988 E-6	004320	.002045	2947
b/c ^B					5,584	-14.95	53,330
peak	•					69.47	199.0
SC							2,667,000

A - 305-day milk.

B - time of peak yield. C - S is c^{-(b+1)}, adjusted for a.

s	.04264	.02441	.01474	00956	.01236	.01519	.038
Peak	.13793	02397	.08005	00029	06954	.15	
b/c	009181	.02170	02356	04992	.071		
υ	.00061	05908	.08517	.15			:
Ą	.04797	06990	.093				
đ	03679	.064					
Milk	.16						
	Milk ^A	ŋ	p	υ	b/c ^B	peak	s. C

Table IV.3.8. Standarized genetic variances (diagonals) and covariances (off-diagonals).

A - 305-day milk yield.

B - time of peak yield. C - S $is^{c-(b+1)}$, is adjusted for a.

		plicitory pac		ogetn\ c			. / of t- dtagonato / .
	M11k	σ	р	υ	b/c	Peak	S
Milk ^A	1	.171	. 071	107	.079	.849	.055
ß		1	. 870	604	270	.123	012
þ			1	.792	.192	.176	.116
U				1	042	.174	128
b/c ^B					1	024	.437
peak						1	.037
SC							1

Standardized phenotypic variances (diagonals) and covariances (off-diagonals). Table IV.3.9.

A - 305-day milk yield.

B - time of peak yield. C - S is c^{-(b+1)}, adjusted for a.

IV.3.2 Heritabilities, genetic and phenotypic correlations

Computed heritabilities for 305-day milk yield and the lactation curve parameters are in Table IV.3.10. The heritabilities and their standard errors are on the diagonals with the genetic correlations on the upper off-diagonals and the phenotype correlations on the lower off-diagonals.

The heritability for 305-day milk (.16) is lower than reported by Shanks (1979). Heritabilities are low for all curve parameters except for c and peak yield. The heritability for S adjusted for <u>a</u>, is .038 compared to .034 when not adjusted for <u>a</u>. Table IV.3.11, contains the heritability values for the curve parameters for the present study and those reported in the literature. Keeping in mind the models are not the same. Then, the parameters in different models represent different traits. Schneeberger's (1981) heritability values for 305-day milk, <u>a</u>, b and c are all greater than found in the present study. Values in the present study are comparable to those of Shanks (1979) for <u>a</u>, b, c and S, while Shanks values for peak (.15 vs. .23) and 305-day milk (.16 vs. .27) were higher.

Genetically, 305-day milk was positively correlated to b, c, peak yield and S and negatively correlated to <u>a</u> and time of peak yield. These correlations suggest that selecting for 305-day milk should decrease <u>a</u>, increase b, but little change in c would be expected because it has a low genetic correlation with 305-day milk (.004).

Heritabilities $^{\mathrm{A}}$, genetic and phenotypic correlations for	305-day milk yield and lactation curve parameters.
able IV.3.10.	

	Milk	σ	Ą	U	b/c	Peak	S
Milk ^B	.16(.167)	367(1.13)	.397(.761)	.00391(.627)	0869(.990)	.911(.146)	.552(1.44)
a	.170(.353)	.064(.466)	906(.455)	607(.677)	. 322(2,58)	248(1.07)	.495(1.69)
Ą	.0713(.096)	870 ^E	.093(.236)	.726(.417)	290(2.66)	.687(.794)	.248(.292)
U	107(.411)	604(.126)	.792 ^E	.15(.164)	487(2.14)	00213(.573)	184(3.73)
b/c ^C	.0788(.548)	270(.941)	.192(.582)	0417(.058)	.071(.234)	683(.172)	.238(.554)
Peak	.849 ^E	.123(.352)	.176(.141)	.174(3.08)	0237(.027)	.15(.164)	.204(1.28)
s ^D	.0551(.100)	0117(.035)	.116(.248)	128(.560)	.437(1.76)	.0366(.155)	.038(.464)

A - The diagonals are heritabilities, the upper off-diagonals are genetic correlations, the lower off-diagonals are phenotypic correlations and standard errors are in parenthesis.

B - 305-day milk yield.

C - Time of peak yield

D - S is $c^{-(b+1)}$, adjusted for a.

E - Variance estimates computed using the formula presented by Grossman (1970) were negative.

Schneeberger (1981) found negative correlations between 305-day milk and b (-.09) and milk and c (-.14) and a positive correlation (.37) with <u>a</u>. The measures of persistency Schneeberger used were positively correlated to 305-day yield (.07 to .16).

The phenotypic correlations between 305-day milk and the curve parameters in the present study were all low except for peak yield (.849) (Table IV.3.10). Phenotypic correlations between <u>a</u> and b (-.87), and <u>a</u> and c (-.604) were negative. The phenotypic correlation between b and c (.792) was positive.

Variables or parameters	Present Study	Shanks (1980)	Schneeberger (1981)
ln a	.06 ^A	.10	.09
Ъ	.06	.06	.15
с	.09	.14	.20
b/c time of peak	.15	.02	
peak	.15	.23	
S	.04	.02	
305-day milk	.16	.27	42

Table IV.3.11 Heritability values reported for lactation curve parameters and 305-day milk.

A - model used computed a, not ln(a).

The genetic correlation between <u>a</u> and b (-.906) and <u>a</u> and c (-.607) were negative and between b and c (.726) was positive. The parameter c is itself a measure of persistency because it

represents the slope after peak yield and persistency may be defined as the ability to maintain peak production. To increase persistency, c needs to be decreased. Therefore, selecting negatively for c to increase persistency would tend to increase <u>a</u> and decrease b both genetically and phenotypically. Shanks (1979), using the logarithm form of Wood's (1967) equation, found very low positive genetic correlations for $\ln(a)$ with b and $\ln(a)$ with c. This differed with the present study and with results of Schneeberger (1981) who found high negative genetic correlations between <u>a</u> and b (-.79) and a and c (-.67).

Based on the genetic correlations, one would expect that selecting to increase <u>a</u> will decrease 305-day milk and b, and decrease c, which would increase persistency. It would also increase S, persistency as Wood (1967) defined it. Selecting to increase b would increase milk, decrease <u>a</u> and increase c. Selecting negatively on c in order to increase persistency would increase <u>a</u> and decrease b and increase S.

Since lactation persistency increases as c decreased it would be expected to be negatively correlated to Wood's (1967) definition for persistency. Although this correlation (-.184) is negative, it is low. Also, those variables positively correlated to c would be expected to be negatively correlated to S and vice-versa. This is not true for b, because b has a positive correlation with both c (.726) and S (.248).

IV.4 Genetic and Correlated Genetic Change

Genetic change was computed for 305-day milk yield selected alone. This and the correlated genetic responses in the lactation curve parameters after one round of selection are reported in Table IV.4.1. The genetic change in milk was 359 lbs. With this, an estimated change of -.475 in <u>a</u>, .006589 in b, 8.678 10^{-7} in c and an increase of 1.15 lbs in the peak yield would be expected. The change in the curve parameters when selecting for milk alone is also expressed relative to their means (Table IV.4.1). The change in the curve parameters when selecting for milk alone was compared to expected change when selecting for each parameters alone, Table IV.4.1. By selecting for milk alone the percent change in <u>a</u>, relative to selecting for <u>a</u> alone, was -160, 52% for b, .4% for c, -113 for time of peak, 95% for peak yield and 113% for S.

The genetic change when selecting alone for each curve parameter is reported in Table IV.4.2, along with the change relative to their means. The genetic change as a percent of the means when each trait was selected alone were 2.4% for 305-day milk 2.6% for <u>a</u>, 5.9% for b, 6.9% for c, 7.7% for time of peak and 2.1% for peak. Table IV.4.3 contains the correlated genetic change in 305-day milk when selecting for each curve parameter alone. The greatest lose in 305-day milk occurs when selecting for <u>a</u> alone (-120%), while the greatest gain occurs when selecting for peak alone (87%).

Table IV.4.4 contains the change in the other curve parameters when selecting for either <u>a</u>, b, or c alone. When selecting for <u>a</u>, the mean changes in b and c were -4.5% and -2.8%, respectively.

Selecting for b alone resulted in a mean shange of -2.8% and 4% in <u>a</u> and c, respectively, and selecting for c alone resulted in a mean change of 2.4\% for <u>a</u> and -5.5% for b. These indexes are compared to genetic gain when selecting for milk alone. Table IV.4.5 lists the indexes with the genetic change for each parameter when that index is used. Table IV.4.6 lists the indexes with the percent genetic change for each parameter to the change when selecting for milk alone. Table IV.4.6 lists the indexes with the percent when selecting for milk alone. Table IV.4.6 lists the indexes with the percent when selecting for milk alone. Table IV.4.6 lists the indexes with the percent genetic change for each parameter relative to the change when selecting for the parameter alone.

Table IV.4.1 Change in lactation curve parameters after one generation of selection for 305-day milk yield alone.

Variables or parameters	unit change	% change 1 ^A	% change 2 ^B	
305-day milk	359.0	. 2.4	100	
a	475	-1.5	-160	
Ъ	.006589	3.1	52.2	
c	8.678 E-7	.021	.411	
b/c time of peak	6950	-1.0	-113	
Peak yield	1.159	2.0	95.3	
s ^C	70.62		114	

- A Change 1 is percent change relative to mean of parameters.
- B Change 2 is percent change relative to change when selecting for parameter alone.
- $C c^{-(b+1)}$, adjusted for a.
- D Mean for S is zero because values are residuals of regression on a.

 Variables or parameters	Unit change	% change ^A
а	.8196	2.6
b	.01262	5.9
С	.0002087	6.9
b/c time of peak	5.306	7.7
Peak yield	1.217	2.1
s ^B	62.06	c

Table IV.4.2	Change in each	lactation	curve parameter
	when selecting	alone for	itself.

A - relative to mean of parameter.

B - S is $c^{-(b+1)}$, adjusted for a.

C - mean for S is zero because values are residuals of regression on a.

Table IV.4.3 Change in 305-day milk yield when selecting for lactation curve parameters alone.

Parameter being selected	Change in milk	% change ^A
' a	-83.4	-120
Ъ	108	30
с	1.38	.38
b/c time of peak	-20.7	-106
Peak yield	313	87
s ^B	96.0	27

A - change in 305-day milk yield relative to its mean. B - S is $c^{-(b+1)}$, adjusted for a.

Parameter	Unit Change	% Change 1 ^A	% Change 2 ^B
When select	ting for a		
a	.8196	2.6	-273
Ъ	009522	-4.5	-244
С	00008339	-2.8	-9700
Peak yield	1999	34	-177
When select	ting for b		
a	8920	-2.8	47.5
b	.01262	5.9	191
с	.0001198	4.0	13800
Peak yield	.6651	1.1	57.4
When select	ting for -c	······	
a	.7557	2.4	-259
Ъ	01159	-5.5	-276
С	0002087	-6.9	-24100
Peak yield	.002448	0.0	0.211

Table IV.4.4 Change in lactation curve parameters when selecting for a, b or c alone.

A - Change relative to mean of parameter.

.

B - Change relative to change when selecting for milk alone.

					for vario	us indexe	s after 1 gen	eration of	selection.	
Index		Var	[ab]	les/We	eights		Genetic Ch	ange of Var	iables	
	MIIk			٩	Peak	305 ^A	. 03	م	υ	PeakA
					Α	ay M11k				
1:1	3			1	2	352	7019	.009087	4.193 E-5	1.315
1:2	'n			٦	1	358	7066	.01167	4.122 E-5	1.324
1:3	ŝ			٦	1	358	6900	.01086	3.740 E-5	1.295
1:4	Г			-	9	340	7016	.01299	2.121 E-5	1.326
1:5				9	9	311	7240	.01450	5.835 E-5	1.302
1:6				7	15	341	6953	.01294	4.547 E-5	1.326
	Milk	в	م	υ	Peak	_				
1:7	1	0	m	0	9	107	008383	.009869	-3.377 E-6	1.246
1:8		0	9	0	و	104	008406	.01073	-4.138 E-6	1.210
1:9	Ś	0	Г	0	п	228	.001455	.006549	-3.595 E-6	.7523
	FW	1k		Peak						
1:10		F1		9		340	3860	.01590	-2.093 E-6	1.227
1:11		e		٦		352	2597	.007731	-2.377 E-7	1.188
1:12		9		Ч		352	4473	.01311	-8.160 E-8	1.177
	FW	1k		٩						
1:13				9		281	9702	.01363	9.585 E-5	1.160
1:14		Г		15		259	9802	.01379	1.022 E-4	1.105
1:15		e		I		358	7468	.01042	4.420 E-5	1.271
1:16	-	9		I		359	7043	.009816	3.690 E-5	1.260

Table IV.4.5 Genetic change in 305-day milk and curve parameters

					s	63.60	1 97.90	93.80	5 92.14		33.46	3.493	-5.464		61.96 1 22.29
		Peak	1.149	.3847	Peak	5530	.09031	1.097	08345		4270	.5405	1120		.8065 .01873
	bles	U	671 E-7	663 E-7 151 E-7	b/c	9.515	7.568	.7042	8.233		7.007	-4.614	-1.482		
	/aria		9	- 2		E-5	E-6	E-6	E-6		E-4	E-4	E-4		E-4 E-4
	nge of V	٩	003771	516 E-4 202 E-5	ပ	-1.937	-5.274	9.635	-8.416		-1.873	1.688	1.540		-1.343 -2.082
	etic Chan		330 .(865 <u>1</u> .	م	009615	005218	.004866	006515		01230	.01500	.01030		3.007 E-3 01161
	Gen	63	900 -	.005	æ	- 08387	2338	4276	1856		. 6949	8054	6004		. 1638 -
		305 Day Milk	102	224 225	305 Day Milk	-7.1	157	356	115		-37.6	176	39		243.9 -3.1
	ghts	U	Õ	00	s		9		10	Peak		10	Ч		
con't)	les/Wei	۹	9		Peak	1		Ч	1	b/c	15	10	9	C	-10
4.5 (ariab	Ø	0	00	b/c	15	Ś	e	10	υ	-10	-10	-9		
ole IV.	Λ	Milk		m vo	Milk	1	-	2		Milk		-1		MIIM	
Tal	Index		1:17	1:18 1:19		2:1	2:2	2:3	2:4		2:5	2:6	2:7		2:8 2:9

Table I	V.4.5	(con'	t)				:				
Index		Varia	bles/	Weights			Genet	itc Change	of Variable	S	
	Milk		Ą	Peak	S	305 Day Milk	cj	Ą	U	Peak	S
3:1	1	10			10	Q	.4222	00313	-2.639 E-6	.3174	137.0
3:2	Ч	10	-5	-5	10	-243	.8437	01410	-5.095 E-5	8604	55.15
3:3	цц.	- 10	-10 -	-10	10	-282	.8324	01520	-5.773 E-5	-1.092	18.62
	Milk	CD	٩	υ	Peak						
3:4		10		-10	ы	-16.9	.9679	004453	-2.832 E-4	.7262	15.30
3:5	1	10	1	-10	ŝ	-171	1.112	01161	-2.296 E-4	2932	-22.60
3:6		10	-10	-10	-10	-235	1.015	01396	-1.602 E-4	8498	-41.40

A - 305-day milk and peak yield are in lbs.

		Peak	110	114	112	114	112	114		107	104	65		106	101
rs in alone	to change lk alone	υ	4830	4750	4310	2440	6720	5240		-489	-577	-514		-341	-109
rve paramete ing for milk	nge relative cting for mi	Ą	138	253	164	197	220	196		149	162	66		241	111/
milk and cu when select	Percent cha when sele	æ	149	148	145	148	157	146		1	1	-100		81 55	60 46
for 305-day ve to change		305 Day Milk	98	100	100	95	87	95		30	29	64		95	9 8 8
change relati n.	ts	Peak	2	1	1	9	9	15	Peak	9	9				
genetic indexes neratio	s/Weigh								υ	0	0	0	Peak	9-	
rcent ríous r l ge	riable	Ą		Г	Ч		9	-1	Ą	e	9	Ч			
IV.4.6 Pe va fo	Va	Milk	m	ິຕ	Ś		-	-	Milk a	1 0	1 0	5	WIIW	6	9
Table	Index		1:1	1:2	1:3	1:4	1:5	1:6		1:7	1:8	1:9		1:10	1:12

	eak	100 95 109		128 33	28	S	90 138	133	130		47 5 -108
	- d					Peak	-148	95	-107		-137 47 -110
nge e	υ	11000 11800 5100		76 -130	-136	b/c	-1470	-201	-1284		-1110 663 213
ve to cha milk alon	٩	207 209 158		57 2	-101	ບ	-2330 -708	1100	-1100		-21700 19500 17700
e relati Ing for		+ 10 ~ 0				٩	-246	4L	-198		-286 228 156
it change selecti		206 206 151	4	-10]	-10	ŋ	18 ,0	66	36		-246 169 126
Percen when	305 Day Milk	78 72 100		29 62	63	305 Day Milk	-102	66	32		-110 49 11
hts	eak	15 15 1	U	00	0	s	-1 v	-	10	Peak	10 10
es/Weig	Ъ		م	9 1 9		Peak	, F-1 F-		1	b/c	15 10 6
Variabl	í 1 k	н н е м	b ct	00	0	b/c	15 4	יא כ	10	υ	-10 -10 -6
	W		MIIk		9	MIIk				MIIK	
Index		1:13 1:14 1:15) - -	1:17 1:18	1:19		2:1	2:3	2:4		2:5 2:6 2:7

Table IV.4.6 (con't)

			S	194 78	26		21 -132	-158
change alone	10		Peak	27 -174	-198		63 -125	-173
lative to o for milk a	0	600 81 000 31	υ	-404 -5970	-6750		-32700 -26600	-18600
change re selecting	Ą	-145 -15(-276 -25(Ą	-147 -313	-330		-168 -276	-311
Percent when	ta	-134 -259	G	-188 -277	-275		-334	-314
	305 Day Milk	68 -101	305 Dav Milk	2 -168	-179		-105 -148	-165
\/Weights	U I	-1 -10	eak S	1 10 -5 10	10 10	c Peak	-10 1	-10 -10
Variables	Milk	1	Milk a b F	1 10 1 1 10 -5	1 10 -10 -	Milk a b		1 10 -10
Index		2:8 2:9		3:1 3:2	3:3		3:5	3:6

Table IV.4.6 (con't)

	Peak	108	109	106	109	107	109		102	66	62		100	96
when milk or	с ^А	-120	-119	-118	-110	-128	-122		1.6	2	1.7		1	.1
to change cted alone	م	72	92	86	103	115	102		78	85	52		126	61 104
ge relative ter is sele	ŋ	-185	-186	-184	-185	-188	-185		-101	-101			-147	-132 -154
Percent chan parame	305 Jay Milk	98	100	100	95	87	95		30	29	64		95	98 98
ghts	Peak I	2	-	-1	9	9	15	c Peak	9 0	0 6	0	Peak	ų	. – –
ables/Wei	٩	1	1	1	Ч	9	1	Ą	ñ	9	-1			,
Vari								B	0	0	0			
	ILIM	9	ĥ	ŝ	-	Ч		Milk	н	Ч	Ś	11 IM	-	e ni
Index		1:1	1:2	1:3	1:4	1:5	1:6		1:7	1:8	1:9		1:10	1:11

Percent genetic change in 305-day milk and curve parameters relative to change when milk or parameter is selected alone for 1 generation. Table IV.4.7.

	Peak	95 90 104 104	94 32 27	Peak S	-145 102 7 158 90 151 -107 148	135 54 44 5 109 -109
change when alone	U	-146 -149 -121 -128	-100 .1	b/c I	179 143 13 155 -	132 -1 -187 -1 -128 -1
lative to s selected	Ą	108 109 82 78	29 1.2 -101	υ	9 2.5 -105 4	-181 -174
ange re meter 1	a	18 19 82	01	P	-176 -141 39 -152	-197 119 81
rcent ch para			1	g	-110 -128 -152 -123	85 -198 -173
Per	305 Day Milk	78 72 100 100	29 63 63	305 Day Milk	-102 44 32	-110 49 11
ghts	۹.	6 15 1 1	v 000	k S	10 1 6 1	Peak 1 10 1
ss/We1			н с д	Peal		b/c 15 10 6
riable			a 000	b/c	15 6 10	-10 -10 -10
Va	MIIk	0 01111	Mflk 1 3 6	MIIk	нчрч	Milk 1 1
Index		1:13 1:14 1:15 1:16	1:17 1:18 1:19		2:1 2:2 2:3	2:5 2:6 2:7

Table IV.4.7 (con't)

Index	2	/ariab	les/	Weight	S	Percen	t chang sel	ge of para ected alon	neters wi e	nen	
		M	k	U		305 Day Milk	đ	م	υ	S	
2:8 2:9				-10		68 -101	20 92	-124 -192	64 100	100 36	
	M11k	60	٩	Peak	s	305 Day Milk	63	Ą	υ	Peak	S
3:1 3:2		10	-1 'r	-	10 10	2 -168	51 103	-125 -212	1 24	26 -171	220 88
3:3		10 -	.10	-10	10	-179	102	-202	28	-190	30
	MIIK	ta	٩	IJ	Peak						
3:4 3:5		10	 - 2	-10 -10	- 2 - 2	-105 -148	118 135	-135 -192	135 110	59 -124	25 -136
3:6		10 -	-10	-10	-10	-165	123	-210	76	-170	-166
A - Para	meter	c, wh	ien s	electe	d alo	ne, was sele	cted no	egatively,	therefo	re a positive	percent

perce	
positive	ln c.
8	e T
therefore	tive chang
ly,	-88a
lve.	ŭ
gati	ŝ
gan	sent
ed	sere
ect	rep
sel	ne
as	alo
3	ບ ·
one	for
al	n8
ted	cti
lec	ele
86	S
rhen	re t
ي م د ا	ttv
л С	cele
letε	ie I
Iran	ang
Pa	Ŀ
1	

Table IV.4.7 (con't)

Because none of the lactation curve parameters that have a high correlation with 305-day milk, has a heritability higher than that of milk, it is expected that selecting for milk alone will increase 305-day milk the fastest. This is evident by the results in Table IV.4.5 where genetic gain in milk is less than 359 lbs for all but a few indexes.

Indexes 1:1 through 1:19 attempt to increase the ascent to peak and/or increase the peak yield while selecting for 305-day milk. The notation 1:1 refers to the first strategy and the first index within that strategy. The greatest increases in milk yield occurred in indexes where milk had weight of 3 or greater, except for index 1:9 table IV.4.5 Index 1:9 is a restricted index, along with 1:7 and 1:8. These restricted indexes restricted the change in a and c while selecting for b. This was done to increase peak yield, $a(b/c)^{b}exp(-b)$. However, this also restricted the gain in any correlated traits, even when they were selected for, i.e., milk, b and peak yield. Kempthorne and Nordskog (1959) discussed restricted index method where $b = P^{-1}Ga$ becomes $b = [I-P^{-1}GC(C'CP^{-1}GC)^{-1}C'G]P^{-1}Ga$ and C'G is the rows of the genetic variance-covariance matrix for the traits being restricted.

The greatest increase in b occurred using indexes 1:5, 1:6, 1:10, 1:12, 1:13 and 1:14 (Table IV.4.5). Index 1:10 increased b, the ascent to the peak, the greatest, although b was not included in the index. This index selected for milk and peak with weights of 1 and 6, respectively.

The largest increase in peak of this first group of indexes occurred in indexes 1:4 and 1:6. The least occurred in two restricted indexes, 1:18 and 1:19, where <u>a</u> and <u>c</u> were restricted. In 1:17, <u>a</u> and <u>c</u> were also restricted, but by putting more selection pressure on b, peak yield increased much more than in 1:18 and 1:19.

The effect of this first group of indexes on c varied from an increase of 9.58 10^{-5} in 1:13 to a more desirable change of -4.13 10^{-6} in 1:8. However, 1:8 restricts c and c only changed 2% of what it would decrease if it were selected alone (Table IV.4.7). However, c changes -577% of that when selecting for milk alone, (Table IV.4.6), but remembering the genetic correlation is only .004 between 305-day milk and c.

In summary, the use of indexes for strategy 1, in which the attempt was to improve b and peak yield, resulted in:

- The greatest change in b, relative to selecting for milk alone occurred in 1:3 and 1:5, and for peak occurred in 1:2, 1:4, 1:6 and 1:17.
- 2) Index 1:2 provided the greatest increase in b and peak while maintaining the same change in milk, as when selecting for milk alone.

The second group of indexes, 2:1 through 2:9 attempted to delay the time of peak and increase persistency. Selecting for persistency was either done by selecting negatively for c or selecting positively for S, Wood's (1967) definition for persistency, $c^{-(b+1)}$, but adjusted for the scaling parameter, <u>a</u>. Only one of these indexes, 2:3, maintained genetic gain for milk production. This index used weights of 7, 3, 1 and 1 for milk, time of peak, peak and S, respectively. Several indexes produced negative gains in 305-day milk; 2:1, 2:5 and 2:9.

In Wood's (1967) equation, selecting for b increases the time to peak. However, for these indexes, time of peak (b/c), is selected for directly. Indexes 2:1 and 2:4 produced the greatest increases in the time to peak. Both of these indexes had negative changes in b and c. To get a larger value for b/c, c must decrease faster than b.

S increased the most in 2:2, but nearly as much in 2:3, when milk and time of peak had higher weightings. Indexes 2:5, 2:8 and 2:9 decreased c the greatest, but this decrease, which indicates an increase in persistency was not consistent with Wood's (1967) measure of persistency, S, which only increased mildly due to use of these indexes (2:5, 2:8, 2:9).

Peak yield did not increase in all indexes, in the second strategy, in which it was selected. It had mild increases compared to the increases in the first group of indexes (1:1 to 1:19). In fact, the greatest increase in peak in the second group occurred when milk was selected the heaviest (2:3).

Most of the group 2 indexes produced negative gains in b. Selecting for b alone would increase time of peak, due to the relation of b/c. However, selecting positively for b was not included in these indexes to delay time of peak. The greatest

decrease in b occured in 2:5, when c was selected negatively. A large increase in b/c resulted (7 days). Thus, both b and c decreased but c is decreased faster.

Comparing indexes 2:8 and 2:9 indicates that selecting negatively for c is detrimental to 305-day milk. This is true even when c is equally weighted with milk (2:8). When c recieves a weight 10 times milk (2.9), the genetic change in milk becomes negative. But the greatest change in c occurs in this index.

Index 2:5 gives the greatest decrease in b and c relative to the genetic change expected when they are selected alone (Table IV.4.7), while increasing time to peak 132% of that when selected alone. The greatest change in b/c relative to selecting for it alone occurred in 2:1 (179%). Likewise the greatest change in b/c relative to when selecting for milk alone occurred in 2:1 (1369%).

The third group of indexes, 3:1 through 3:6, attempts to flatten the lactation curve by increasing <u>a</u>, decreasing the peak and increasing persistency or decreasing c. Indexes 3:1 to 3:6 (Table IV.4.5) indicate this is not possible without decreasing milk yield considerably. The greatest loss in milk occurred in 3:2, 3:3 and 3:6.

The increase in <u>a</u> was greatest in 3:5 and 3:6. All indexes used caused a decrease in b, with the greatest decrease in 3:2 and 3:3, in which b was being selected negatively.

Selecting for S was most successful in 3:1 and the only index where the genetic change in 305-day milk was not negative. This index produced a 220% increase in S, relative to selecting for it alone (Table IV.4.7). Indexes 3:4 and 3:5 where most successful in selecting negatively for c. In fact, the most successful in all three groups of indexes. Index 3:4 and 3:5 produced greater decreases in c than selecting negatively for it alone, -135 and -110%, respectively.

Indexes 3:4, 3:5 and 3:6 produced greater increases in <u>a</u> than selecting for it alone, 118, 135 and 123%, respectively. Concurrently, 3:4 and 3:5 produced the most desired results for <u>a</u> and c based on the goals of this group of indexes. The percent change in <u>a</u> and c relative to selecting for them alone was 118 and -135 for 3:4 and 135 and -110 for 3.5, respectively.

IV.4.1 Changes in lactation production

The genetic change in 305-day milk and peak yield were computed in two ways. 1) From the expected genetic change in 305-day milk and peak for each index. These will be called the expected values. 2) From the expected genetic change in \underline{a} , b and c, estimates for 305-day milk and peak yield were computed from:

 $\hat{y}_{305-\text{day milk}} = a_0 \int^{305} t^b \exp(-ct) dt$ and

Peak = $a(b/c)^{b}exp(-b)$.

These values will be called estimated values.

For selected indexes, these expected and estimated values are reported in Tables IV.4.8 to IV.4.10 for 1, 5 and 10 generations of selection. In the same tables, change between generations and the accumulated change from generation to generation for expected and estimated 305-day milk are reported. The expected values for the curve parameters <u>a</u>, b and c are in Table IV.4.11 to IV.4.13 for generations 1, 5 and 10.

The base values, which are those computed for the current population, are 16,684 lbs for 305-day milk, 62.9 lbs for peak, 70.2 days for time to peak, 31.6 for <u>a</u>, .212 for b and .00302 for c. These values are listed in Table IV.4.11 as generation zero in the milk only index. The value for 305-day milk, 16,684 lbs, is the estimated value from the integral produced by by base values for <u>a</u>, b and c. These base values are the population means prior to adjustment for age and are reported in Table IV.3.3.

When selecting for milk only (first index, Table IV.4.8 and IV.4.11), the expected 305-day milk yield was 17,043, 18,479 and 20,274 lbs for generations 1, 5 and 10, respectively. However, this was not equal to the estimated change from computing new curves when selecting for milk alone. They were 16,949, 18,011 and 19,329 lbs, respectively for 1, 5 and 10 generations. The expected values were about 100 lbs more per year than those values estimated by the integrals (Table IV.4.14). Comparison of the expected and estimated values for peak yield also shows the estimated values were less than the expected values (Table IV.4.14).

Index/Gen	Vaț	riabl	es/Weigh	ts Exp	ected		Estimated				
				305 Daý Milk	Accum ^A Change Milk	305 Day M11k	Change ^B Milk	Accum ^C Change Milk	Est. Peak	Expected Peak	Expected b/c
0/	Base	e Gen	eration	16684	ł	16684	1	· 1	62.9	62.9	70.2
	4LtM										
M11k/1	Н			17043	359	16949	264	264	63.7	64.0	72.3
S				18479	1745	18011	1061	1326	67.1	68.7	80.9
10				20274	3590	19329	1318	2640	71.3	74.5	91.7
	Milk	p	Peak								
1:1/1	e	F	2	17036	352	16922	237	237	63.8	64.2	72.2
c 10				18444 20204	1/60 3520	1/816 18738	89.5 922	2053	0. 4 70. 4	c.60 0.97	/9./ 88.0
	M11k	٩	Peak								
1:4/1	1		6	17024	340	17289	604	604	64.9	64.2	73.9
5				18384	1700	19846	2556	3161	73.7	69.5	88.5
10				20084	3400	23292	3446	6607	85.9	76.2	105.7

5 and 10 genetic values for 305-day milk, peak and time Generations 1, Table IV.4.8.

Index/Gen	Variables/Weight	EX	pected		Estimate	q			
	Milk a b c Pea	305 ik Day Milk	Accum ^A Change Milk	305 Day Milk	Change ^B Milk	Accum ^C Change Milk	Est. Peak	Expected Peak	Expected b/c
1:7/1 5 10	1 0 3 0 6	16791 17219 17754	107 535 1070	17482 21087 26703	797 3605 5616	797 4402 10018	65.9 78.1 97.7	64.1 69.1 75.4	73.5 87.0 103.9
	Milk a b c Peak	1.4							
1:8/1 5 10	1 0 6 0 6	- 16788 17204 17724	104 520 1040	17554 21533 27856	869 3979 6322	869 4848 11171	65.9 79.6 101.8	64.1 68.9 75.0	73.8 88.5 107.1
	Milk a b c Peak	1.4							
1:9/1 5 10	5 0 1 0 1	- 16912 17824 18964	228 1140 2280	17216 19526 22873	531 2309 3346	531 2841 6188	64.7 72.6 84.2	63.6 66.7 70.4	72.4 81.5 92.9
	Milk Peak					·			
1:10/1 5 10	1 6	17024 18384 20084	340 1700 3400	17767 22857 31328	1082 5089 8491	1082 6172 14643	66.5 84.0 114.3	64.1 69.0 75.2	75.4 96.8 123.6

Table IV.4.8 (con't)

Index/Gen	Variable	ss/Weight	s Expec	sted	I	Istimated				
			}							
			305	Accum	305	æ	Accum			
	MIIK	Peak	Day Milk	Change Milk	Day Milk	Change ⁷ Milk	Change Milk	Est. Peak	Expected Peak	Expected b/c
1:11/1	e e		17036	352	17161	476	476	64.5	64.1	72.7
S			18444	1760	19202	2041	2517	71.4	68.8	83.0
10			20204	3520	22085	2882	5400	81.3	74.8	95.8
	MIIK	p								
1:13/1		6	16965	281	17004	319	319	64.2	64.1	72.4
ŝ			18089	1405	18153	1149	1468	68.9	68.7	80.0
10			19494	2810	19116	963	2431	73.3	74.5	87.5
	MIIK	٩	·							
1:16/1	9	1	17044	360	16691	307	307	64.0	64.2	72.5
ŝ			18484	1800	18185	1193	1500	68.2	69.2	81.4
10			20284	3600	19525	1340	2840	73.0	75.5	91.5
	Milka	. .c								
		, ,	_							
1:17/1	1 0	6 0	16786	102	16978	293	293	63.9	64.0	71.4
2			17194	510	18206	1228	1522	68.1	68.6	76.3
10			17704	1020	19873	1666	3188	73.9	74.4	82.5

Table IV.4.8 (con't)

.

Index/Gen	Vari	antes	s/Wei	ghts	Expe	cted	E	stimated				
	M11k	57	م	U	305 Day Milk	Accum ^A Change Mi1k	305 Day Milk	Change ^B M11k	Accum Change Milk	Ext. Peak	Expected Peak	Expected b/c
1:19/1	9	0	-	0	16909	225	16683	-1.5	-1.5	62.9	63.2	70.2
Ś					17809	1125	16677	-6	-7.5	62.9	64.5	70.1
10					18934	2250	16670	-7.4	-14.9	62.9	66.2	70.0

Table IV.4.8 (con't)

B - Change between generations for estimated 305-day milk.

C - Accumulated change from base generation for estimated 305-day milk.

Index/Gen	Vari	ables/	/Weigh	ts	Exp	ected		Estimate	q				
					305 Day Milk	Accum ^A Change Milk	305 Day Milk	Change ^B Milk	Accum ^C Change Milk	Est. Peak	Expected Peak	Est. b/c	Expected b/c
2:1/0	Ba	se Gen	lerati	uo	16684	1	16684	ł	1	62.9	62.9	70.2	70.2
	Milk	b/c	Peak	S									
2:1/1 5 10		15			- 16677 16649 16614	-7 -35 -70	15950 13328 10660	-734 -2622 -2668	-734 -3356 -6024	60.3 51.2 42.1	62.4 60.2 57.4	67.4 56.0 41.0	79.7 117.7 165.4
	MIIk	b/c	Peak	လု									
2:4/1 5 10	-	10	-	10	16799 17259 17834	115 575 1150	16106 13985 11719	-578 -2121 -2265	-578 -2699 -4964	60.8 53.5 45.6	62.8 62.5 62.1	68.2 60.2 50.0	78.4 111. 152.
	Milk	υ	b/c	Peak									
2:5/1 5 10		-10	15		16646 16494 16304	-38 -190 -380	16539 15908 15042	-145 -631 -865	-145 -776 -1642	61.9 57.4 51.9	62.5 60.8 58.7	70.4 72.2 77.5	77.2 105.0 140.0

5 and 10 genetic values for 305-day milk, peak and time Generation 1. Table IV.4.9.

Index/Gen	ı Variab	les/We	1 ghts	Ext	ected		Estimate	q				
	Milk c	b/c	Peak	305 Day Milk	Accum ^A Change Milk	305 Day Milk	Change ^B M11k	Accum ^C Change M11k	Est. Peak	Expected Peak	Est. b/c	Expected b/c
2:6/1 5 10	1 -10	10	10	16860 17564 18444	176 880 1760	17024 18327 19721	339 1303 1394	339 1642 3036	64.6 71.2 78.9	63.5 65.6 68.3	71.2 74.3 76.9	65.6 47.1 24.1
	MIIK	U										
2:9/1 5 10	н	-10		16681 16669 16654	-3 -15 -30	16676 16569 16305	-8.7 -106 -264	-8.7 -115 -379	62.2 59.3 56.5	62.9 63.0 63.1	71.2 77.7 102.	

Table IV.4.9 (con't)

A - Accumulated change from base generation for expected 305-day milk.

B - Change between generations for estimated 305-day milk.

C - Accumulated change from base generation for estimated 305-day milk.

		pea	uk for gi	roup 3 in	dexes.					
Index/Gei	ı Variable	s/Weight	S Ex	bected		Estimate	p			
			305 Day Milk	Accum Change Milk	305 Day Milk	Change ^I Milk	3 Accum ^C 3 Change Milk	Est imated Peak	Expected Peak	Expected b/c
. 0/	Base Gener	ration	16684	1	16884		1	62.9	62.9	1
	Milk a b	Peak S								
3:1/1	1 10 1	1 10	16690	900	16667	-17	-17	62.9	63.2	69.1
د 10			16/14 16744	90 90	16570 16395	-97 -175	-114 -289	62.9 62.7	64.5 66.1	65.2 60.3
	Milk a b	Peak S								
3:3/1	1 10 -10	-10 10	16402	-282	16079	-605	-605	60.8	61.8	66.4
5			15274	-1410	13795	-2283	-2889	53.1	57.5	49.8
10			13864	-2820	11281	-2514	-5402	45.6	52.0	24.6

Index/Ge	ធ្ល	Var	1ab1	es/V	leight	s Expe	cted	Γ	Estimated				
	MII	ک	م	U	Peak	305 Day Milk	Accum Change Milk	305 Day Milk	Change ^B Milk	Accum ^C Change Milk	Estimated Peak	Expected Peak	Expected b/c
3:4/1 5 10		F	0 1	-10		16667 16599 16514	-17 -85 -170	17552 21410 27282	868 3857 5872	868 4725 10597	64.9 74.5 108.6	63.6 66.6 70.2	75.8 118 872
	Milk	æ	٩	U	Peak								
3:6/1 5 10		10.	-10	-10	-10	16449 15509 14334	-235 -1175 -2350	16508 15682 14472	-176 -825 -1210	-176 -1002 -2212	61.9 57.5 51.6	62.1 58.7 54.4	69.2 64.1 51.1
			.	.									

Table IV.4.10 (con't)

A - Accumulated change from base generation for expected 305-day milk.

B - Change between generations for estimated 305-day milk.

C - Accumulated change from base generation for estimated 305-day milk.

Index/Gen	Vari	labl	es/W	eigh	ts				
						a	Ъ	с	
/0	Bas	se G	ener	atio	n	31.6	.21222	.0030243	
	Mil	k							
Milk/1 5 10	1					31.1 29.2 26.8	.21881 .24516 .27811	.0030251 .0030286 .0030329	
	Milk		Ъ]	Peak				
1:1/1 5 10	3		1		2	30.9 28.1 24.5	.22131 .25766 .30309	.0030662 .0032339 .0034437	
	Milk		Ъ]	Peak				
1:4/1 5 10	1		1		6	30.8 28.1 24.5	.22521 .27719 .34217	.0030455 .0031303 .0032360	
	Milk	а	b	с	Peak				
1:7/1 5 10	1	0	3	0	6	31.6 31.5 31.5	.22208 .26156 .31091	.0030210 .0030074 .0029905	
	Milk	a	Ъ	с	Peak				
1:8/1 5 10	1	0	6	0	6	31.6 31.5 31.5	.22295 .26586 .31951	.0030201 .0030036 .0029829	
	Milk	a	Ъ	с	Peak				
1:9/1 5 10	5	0	1	0	1	31.6 31.6 31.6	.21877 .24496 .27771	.0030207 .0030063 .0029883	

Table IV.4.11 Genetic values for curve parameters a, b and c for generations 1, 5 and 10 using group 1 indexes.
Index/Gen	Vari	ables	/Weig	hts			
	M11	.k	Pea	k	a	b	с
1:10/1 5 10	1	_	6		31.2 29.6 27.7	.22819 .29172 .37122	.0030222 .0030138 .0030034
	Mil	.k	Pea	k			
1:11/1 5 10	3		1		31.3 30.3 29.0	.21995 .25088 .28954	.00300240 .0030231 .0030219
	Mi	.lk	Ъ				
1:13/1 5 10	1		6		30.6 26.7 21.9	.22585 .28038 .34854	.0031201 .0035035 .0039827
	Mi	.1k	Ъ				
1:16/1 5 10	6	.	1		30.9 28.0 24.5	.22204 .26130 .31038	.0030612 .0032088 .0033933
	Milk	a	Ъ	с			
1:17/1 5 10	1	0	6	0	31.6 31.5 31.5	.21599 .23107 .24993	.0030249 .0030276 .0030309
	Milk	a	Ъ	с			
1:19/1 5 10	6	0	1	0	31.6 31.6 31.6	.21215 .21186 .21150	.0030239 .0030227 .0030211

Index/Gen	Vari	ables	/Weigh	its			
					a	Ъ	с
/0	B	ase G	Generat	ion	31.6	.21222	.0030243
	Milk	b/c	Peak	S			
2:1/1 5 10	1	15	1	1	31.5 31.1 30.7	.20260 .16414 .11607	.0030049 .0029270 .0028310
	Milk	b/c	Peak	S			
2:4/1 5 10	1	10	1	10	31.4 30.6 29.7	.20570 .17964 .14707	.0030158 .0029820 .0029416
	Milk	с	b/c	Peak			
2:5/1 5 10	1	-10	15	1	32.3 35.0 38.5	.19992 .15070 .089190	.0028374 .0020880 .0011510
	Milk	с	b/c	Peak			
2:6/1 5 10	1	-10	10	10	30.8 27.5 23.5	.22722 .28723 .36225	.0031932 .0038683 .0047131
	M	lilk	с				
2:9/1 5 10		1	-10)	32.3 35.4 39.1	.20061 .15415 .096090	.0028164 .0019833 .00094231

Table IV.4.12 Genetic values for curve parameters a, b and c for generations 1, 5 and 10 using group 2 indexes.

Index/Gen	Variables/Weights			
		а	Ъ	с
/0	Base Generation	31.6	.21222	.0030243
	Milk a b Peak S			
3:1/1 5 10	1 10 1 1 10	32.0 33.7 36.8	.20909 .19656 .18091	.0030216 .0030111 .0029979
	Milk a b Peak S			
3:3/1 5 10	1 10 -10 -10 10	32.4 35.7 39.9	.19702 .13624 .060263	.0029665 .0027356 .0024469
	Milk a b c Peak			
3:4/1 5 10	1 10 1 -10 1	32.5 36.4 41.2	.20777 .18995 .16769	.0027411 .0016081 .00019230
	Milk a b c Peak			
3:6/1 5 10	1 10 -10 -10 -10	32.6 36.6 41.7	.19826 .14242 .072621	.0028641 .0022232 .0014222

Table IV.4.13	Genetic values for curve parameters a, b and
	c for generations 1, 5 and 10 using group 3 indexes.

These discrepancies in 305-day yields when selecting for milk alone indicate that the expected genetic change in 305-day milk based on the equation

$$\Delta G_{milk} = h_{milk}^2 \times \sigma_p \times z/q$$

is not precisely reflected by the change in the shape of the curve generated by the expected values for \underline{a} , b and c over a number of generations.

Discrepancies of this nature occur to greater and lesser degrees in the indexes listed in the Table (III.6.1). Comparison between the expected and estimated accumulative genetic changes in 305-day milk are helpful in seeing the amount of discrepancy that occurs for each index (Tables IV.4.8 to IV.4.10). The discrepancies occur in both directions, i.e., the estimated values both overestimated and underestimated the expected values (Table IV.4.14).

Indexes with the greatest overestimation of the expected values were 1:7, 1:8, 1:10 and 3:4. Index 1:10 overestimated 305-day milk by 10,130 lbs after 10 generations of selection (Table IV.4.14). This means the shape of the curve is much higher and the integral or the area under the curve (31,328 lbs, Table IV.4.8) is much greater than is likely to occur through selection. As mentioned earlier, none of these indexes should yield more 305-day milk than selecting for milk alone. These indexes also overestimated peak yield. For example 1:10 estimated a peak of 114.3 after 10 generations, and the expected value was 75.2 lbs.

	and peak yie	ld for vario	us indexes.				
		30	5-day milk	yield		peak	vield
		Expected					
Index	generation	gain/gen	Expected	Estimated	EST-EXP	Expected	Estimated
M11k	1	359	17043	16949	-94	64.0	63.7
	Ś		18479	18011	-468	68.7	67.1
	10		20274	19329	-945	74.5	71.3
-	1	352	17036	16922	-114	64.2	63.8
3m LD 2 peak	Ŝ		18444	17816	-628	69.5	67.0
T: T	10		20204	18738	-1466	76.0	70.4
•	1	340	17024	17289	265	64.2	64.9
Im ID Opeak	S		18384	19846	1462	69.5	73.7
T:4	10		20084	23292	3208	76.2	85.9
lm Oa 3b Oc	1	107	16791	17483	692	64.2	62.9
6 Peak	S		17219	21087	3868	69.1	78.1
1:7	10		17754	26703	8449	75.4	97.7
1m Oa 6b Oc	1	104	16788	17554	766	64.7	62.9
6 Peak	S		17204	21533	4329	69.0	79.6
1:8	10		17724	27856	10132	75.0	101.8
5m Oa 1b	Ч	228	16912	17216	304	63.7	64.7
Oc 1Peak	5		17812	19526	1702	66.7	72.6
1:9	10		18964	22873	3909	70.4	84.2

Table IV.4.14. Discrepancies between expected and estimated 305-day milk yield

				:				
		(°)	805-day milk	yield		peak >	yield	
		Expected						1
Index E	generation	gain/gen	Expected	Estimated	EST-EXP	Expected	Estimated	
		340	17024	17767	743	64.1	66.5	ł
lm 6Peak	ر ، ا)	18384	22857	4473	69.0	84.0	
1:10	10		20084	31328	11244	75.2	114.3	
	1	352	17036	17161	125	64.1	64.5	
3m 1 Peak	ŝ		18444	19202	758	68.8	71.4	
1:11	10		20204	22085	1881	74.8	81.3	
	Г	281	16965	17004	39	64.1	64.2	
1m 6b	2		18089	18154	-36	68.7	68.9	
1:13	10		19494	19116	-378	74.5	73.3	
	1	360	17044	16991	-53	64.2	64.0	
6m 1b			18484	18185	-299	69.2	68.2	
1:16	10		20284	19525	-759	75.5	73.0	
	1	102	16786	16978	192	64.1	63.9	
1m 0a 6b 0c	5		17194	18206	1012	68.6	68.1	
1:17	10		17704	19873	2169	74.4	73.9	
	1	225	16909	16683	-226	63.2	62.9	
6m 0a 1b 0c	2		17809	16677	-1132	64.5	62.9	
1:19	10		18934	16670	-2264	66.2	62.9	
]m 15bc	1	-7	16677	15950	-727	62.4	60.3	
1 Peak 1S	ŝ		16649	13328	-3321	60.2	51.2	
2:1	10		16614	10660	-5954	57.4	42.1	

Table IV.4.14 (con't)

(con't)
7.4.14 (
Table IV

		305-day п	ilk yield			peak yi	eld
Index	generation	Expected gain/gen	Expected	Estimated	EST-EXP	Expected	Estimated
1	-	115	00731	70171	603	6.7 B	0 09
10.01 100	- v		17750	13005	160- 1765-	07.0 67 5	00.0 52 5
2:4	10		17834	61/11	-32/4	62.1	45.6
1m -10c		-38	16646	16539	-107	62.5	61.9
15bc 1Peak	Ś		16494	15908	-586	60.8	57.4
2:5	10		16304	15042	-1262	58.7	51.9
1m -10c	1	176	16860	17024	164	63.5	64.6
10bc 10Peak	2		17564	18327	763	65.6	71.2
2:6	10		18444	19721	1277	68.3	78.9
	1	6 -	16681	16676	۲ -	62.9	62.2
1m -10c	S		16669	16569	-100	63.0	59.3
2:9	10		16654	16305	-349	63.1	55.5
1m 10a 1b	Г	Q	16690	16667	-23	63.2	62.9
Ipeak 10S	ŝ		16714	16570	-144	64.5	62.9
3:1	10		1.6744	16395	-349	66.1	62.7
lm 10a -10b	Ч	-282	16402	16079	-323	61.8	60.8
-10peak 10S	5		15274	13795	-1479	57.5	53.1
3:3	10		13864	11281	-2583	52.0	45.6
1m 10a 1b	Ч	-16.9	16667	17552	885	63.7	64.9
-10c 1peak	S		16599	21410	4811	66.6	74.5
3:4	10		16514	27282	10768	70.2	108.6

Table IV.4.14 (con't)

		30 5 -d _i	ay milk yile	pa		peak y1	eld
		Expected					
Index	generation	gain/gen	Expected	Estimated	EST-EXP	Expected	Estimated
1m 10a -1 <u>0</u> b	1	-235	16449	16508	59	62.1	61.9
-10c -10peak	ŝ)	15509	15682	173	58.7	57.5
3:6	10		14334	14472	138	54.4	51.6
					•		

.

Indexes which underestimated expected genetic changes in 305-day milk the greatest were 1:19, 2:1, and 2:4 (Table IV.4.8 and IV.4.9). Peak yields were also underestimated in these three indexes. Indexes 1:19 and 2:4 estimated negative changes in 305-day milk, -14.9 and -4,964 lbs by 10 generations when the expected genetic changes were positive, 2,250 and 1,150 lbs, respectively (Table IV.4.8 and IV.4.9). For 2:4, the difference in 305-day milk by 10 generations was 6,115 lbs and the difference in peak yield was 62.1 - 45.6 = 16.5 lbs (Table IV.4.14).

Indexes which had estimates close to the expected 305-day values were 1:13, 1:16 and 3:6. The difference at 10 generations for 3:6 was 760 lbs. Indexes which had estimated peaks close to the expected values where 1:13, 1:16, 1:17, 3:1 and 3:6.

The expected 305-day and peak values were the expected genetic gain times the number of generations. Therefore, the increments between generations were equally spaced. This is not true for the intervals between generations computed from the estimated values. Notable changes in the rates of change from generation to generation occurred in 1:7, 1:8, 1:10, 1:11, 2:9, 3:1 and 3:4 (Tables IV.4.8 to IV.4.10). All of these had an increasing rate of change in estimated 305-day milk from generation 1 to generation 10. The rate of change decreased in 1:13 (Table IV.4.8).

Typically, the rate of genetic change is considered to be constant for a given population and a given selection intensity over a number of generations. This is because h^2 and σ_p do not change greatly after a few generations of selection. Therefore,

$$\hat{\Delta G} = h^2 \mathbf{x} \sigma_{\mathbf{p}} \mathbf{x} \mathbf{z}/q$$

would produce a relatively constant ΔG from one generation to the next. It is therefore disturbing to note that for a constant rate of change of the curve parameters <u>a</u>, b, and c, the rate of change in the shape of the curve is not constant i.e., the rate of change in estimated 305-day yield. In fact, rather disproportionate genetic changes occurred when several indexes were used (1:7, 1:8, 1:10 and 3:4).

The dispropotionate genetic change is estimated 305-daymilk from one generation to the next or the change in the rate of change is due to estimation using a nonlinear equation. When <u>a</u>, b and c change linearly in the equation:

 $\hat{y}_{305-\text{day milk}} = a_0 \int^{305} t^b \exp(-ct) dt$,

then y_{305-day milk}, the total area under the curve, changes nonlinearly. This explains why the estimated change by 10 generations is not 10 times the estimated change in one generation for all indexes (Tables IV.4.8 to IV.4.10). This is also responsible for a small part of the discrepancies between expected and estimated genetic gain in 305-day milk. This is an inherent problem when nonlinear models are used, and one desires to estimate genetic progress with the model.

The computation of genetic correlations assumes linear relationships between traits. Therefore, the relationships between <u>a</u> and 305-day milk, <u>a</u> and b, <u>a</u> and c, b and c, etc., are assumed to be linear. It is possible that some of these relationships are curvilinear as demonstrated in Figure 1. A curvilinear relationship suggests that the correlations change notably as genetic change in the traits occur. This makes it difficult to estimate correlated genetic responses over time. In Figure 1, the correlation would be computed as the best estimate of a linear relationship between x and y. This is represented by the straight line. The linear correlation would only be appropriate within a certain range of x and y.

If the relationship between any of the curve parameters and 305-day milk is nonlinear, then the true correlated genetic response between them would be nonlinear. That is, if 305-day milk and b are nonlinearly related, then as milk changes linearly, b changes curvilinearly or vice-versa. Therefore, when an index is used and curvilinear relationships exist, correlations used for the first generation would not be the same as those used in later generations to compute genetic responses. Therefore, G and P would become dynamic, i.e., contain different covariances over time. It then follows that the b, σ_{I} and Cov (G_{i} , I) become dynamic. Then, in the example of expected 305-day milk, a curvilinear response could be computed.

The correlations would be computed using a polynomial model. Between 305-day milk and <u>a</u> for example, the possibilities may be:

 $y_{305-day milk} = b_1 a + b_2 a^2 + e$

or

$$y_{305-day milk} = b_1 a + b_2 a^2 + b_3 a^3 + e$$

where the change in milk is a polynomial or curvilinear function of \underline{a} .





>

A more complicated situation may exist where the curve parameters <u>a</u>, b and c have curvilinear relationships among themselves. Then a polynomial model is needed for each curvilinear relationship to define the correlations at any level of the parameters.

IV.4.2 Change in the shape of lactation curves

Figures 2 through 4 are plots of the change in the shape of the curves when <u>a</u>, b and c are changed when using the nonlinear form of Wood's (1967) equation. These changes do not consider correlated change in the other curve parameters. Therefore, these curves demonstrate the change due to changing one parameter while holding the other two constant.

Figure 2 shows that as <u>a</u> increased from 25 to 43, the curve maintains its shape, but starts at a higher point. For this reason, <u>a</u> is referred to as the scaling parameter. Figure 3 shows the change in shape as b is increased. The ascent to the peak becomes steeper as b is increased from .24 to .33. Also, the peak and the later stage of the curve increase in height, with the decline after peak, c, remaining constant. Therefore, the area under the curve increases. Figure 4 shows the change in shape as c is increased. The largest value for c (.041) yields the bottom curve with the greatest slope after peak. The most persistent curve is the highest curve which represents the lowest value for c (.032). As c decreases the peak also rises. This is a function of $a(b/c)^{b}exp(-b)$, which also increases as b increases.



Change in the shape of the lactation curve when \underline{a} is changed from 25 to 43 in Wood's equation. Figure 2.







Change in the shape of the lactation curve when c is changed from .032 to .041 in Wood's equation. Figure 4.

DAILY MILK Yield

These curves are a function of t in weeks as Wood (1967) defined time. The values for <u>a</u>, b and c were deviates from those published by Wood (1970), 30, .28 and .036, respectively.

The curves in Figure 5 to Figure 17 are produced using the nonlinear equation $y = at^{b}exp(-ct)$, where t goes from 0 to 305 days. The values for <u>a</u>, b and c are the expected values computed from their correlated genetic change (Table IV.4.5). The discrepancies between the integrals of these curves and the expected 305-day milk need to be kept in mind. Comparisons can be made between the integrals in Tables IV.4.8 through IV.4.10 and the shape of the curves at 1, 5 and 10 generations plotted in these figures.

Figures 5, 6 and 7 show the change in shape at 1, 5 and 10 generations when <u>a</u> is selected alone. The three curves in each figure represent the base (B) or zero generation, the shape when selecting on milk alone (M) and the shape when selecting on the index (I). Again, the base generation was computed using the mean values for <u>a</u>, b and c for the 5,927 first lactation records. As <u>a</u> increases, the index curve has an increase in initial production, but now the peak drops and the slope after peak increases. This is because the correlations between <u>a</u> and b (-.906) and a and c (-.607) were negative. Therefore, as <u>a</u> increases, both b and c decrease, but b decreases faster causing the peak ($a(b/c)^{b}exp(-b)$) and the time of peak (b/c) to decrease. The 305-day milk yield is also negatively correlated with <u>a</u> and therefore some decrease in





Βάίλ **Μίικ Υί**ειά



Change of the shape of the lactation curve after five generations of selection on \underline{a} alone. Figure 6.

Daily Milk Yield



Change of the shape of the lactation curve after ten generations of selection on \underline{a} alone. Figure 7.

the integral would be expected. However, the expected change in 305-day milk was -83, -415 and -830 compared to -116, -666 and -501 for estimated 305-day milk for 1, 5 and 10 generations, respectively (Table IV.4.11).

Figures 8, 9 and 10 graph the change in shape when b is selected alone (I). A slight increase in peak occurs at first and a decrease in time to peak occurs while an increase in c causes an increase in the slope after peak. The greatest change compared to selecting for milk alone is in c, which increases much greater when selecting for b (13,805%) (Table IV.4.5). The correlation between b and c was estimated at .726 while that between 305-day milk and c was insignificant (.004). The expected change in milk is 108, 540 and 1,080 lbs compared to 223, 986 and 1,597 lbs computed by the integral for generations 1, 5 and 10, respectively.

Finally, selecting alone for negative c is represented in Figures 11, 12 and 13. This causes a flattening of the curve (I) and a large loss in production from the base generation (B) even though the expected change in 305-day yield is +1.38 lbs per generation.

Figures 14, 15 and 16 show the curve changes for index 3:4. This is an exmaple of an index in which estimated 305-day milk greatly overestimates expected 305-day milk. The intent of this index was to flatten the shape of the curve by increasing <u>a</u> and decreasing c. The weights were 1, 10, 1, -10 and 1 for milk, <u>a</u>, b, c and peak yield, respectively. Keeping in mind that none











Change in the shape of the lactation curve after ten generations of selection on b alone. Figure 10.

Daily Milk Yield



Change in the shape of the lactation curve after one generation of negative selection on c alone. Figure 11.



Change in the shape of the lactation curve after five generations of negative selection c alone. Figure 12.



Change in the shape of the lactation curve after ten generations of negative selection on c alone. Figure 13.

Daily Milk Yield



Change in the shape of the lactation curve after one generation of selection on index 1 milk, 10a, 1b, -10c and 1 peak yield. Figure 14.

Change in shape of the lactation curves after five generations of selection on index 1 milk, 10a, 1b, -10c and 1 peak yield. Figure 15.



Days in Lactation



Change in the shape of the lactation curve after ten generations of selection on index 1 milk, 10a, 1b, -10c and 1 peak yield. Figure 16.

Daily Milk Yield



Change in shape of the lactation curve after ten generations of selection on index 1 milk and 6 peak yield. Figure 17.

Daily Milk Yield

of the indexes increased expected 305-day milk more than selecting for milk alone. Therefore, none of the index curves (I) should have a greater area (integral) than the curves representing selection for milk alone (M). Noting here that the curves for selection on milk alone (M) represent the 305-day estimates noted in Table IV.4.8 and they underestimate expected 305-day milk slightly. The integral for (I) at 10 generations is 27,282 lbs while that for (M) is 19,329. The expected values for (I) and (M) are 16,514 and 20,274 lbs, respectively. The curve for (I) should therefore be lower than that for (M) (Figure 16). These curves give an idea of the overestimation that occurs when the expected values for <u>a</u>, b and c are used to compute a new index curve (I) for generations 1, 5 and 10 (Figure 14, 15 and 16).

Figure 17 represents another index (1:10) which greatly overestimates expected 305-day milk. This index includes milk and peak yield with weights of 1 and 6, respectively. By 10 generations the (I) curve represents 31,328 lbs and a peak yield of 114 lbs, while the expected values are 20,084 and 76 lbs, respectively. This exceeds the estimated curve for selection on milk only (M) by a staggering amount.

IV.4.3 Summary of changes caused by selection indexes

Without knowing the optimum shape of the lactation curve with regard to efficiency of milk production, one can draw some conclusions about the indexes investigated. If we consider expected values for 305-day milk, we can exclude those indexes

that decrease or slow greatly the genetic progress in 305-day milk yield. It would be unlikely that any change in the shape of the lactation curve which reduces 305-day milk extremely, would produce greater net profit due to reductions in stress and/or inputs. Therefore, of the indexes listed in Tables IV.4.8 through IV.4.10, 1:7, 1:8, 1:17, 2:1, 2:4, 2:5, 2:9, 3:1, 3:3, 3:4 and 3:6 can be excluded. This excludes all but one index (2:6) in group 2 which are attempting to delay peak and/or increase persistency. It also excludes all indexes which attempted to increase <u>a</u> and decrease c or increase S (third strategy). It may be that these indexes would be more desirable if more weight were applied to milk. It was intentional that milk was not selected strongly so that extremes could be compared to selecting for milk alone.

From the values for peak and 305-day production, index 1:13 appears to do a reasonable job of increasing the portion of milk produced in the early part of lactation (Tables IV.4.8). This would be desirable if cows have higher daily net profit in early (peak period) lactation and if this higher production in early lactation is not detrimental to production in subsequent lactations.

Index 2:6 which attempted to delay time to peak, had weights of 1, -10, 10, 10 for milk, c, time of peak and peak, respectively. The expected values for b/c indicate it did not delay time to peak. They were 66, 47 and 24 days for generation 1, 5 and 10, respectively. Indexes 2:1, 2:4 and 2:5 successfully delayed expected time to peak but greatly decreased 305-day milk. The estimated time to peak was actually reduced in 2:1 and 2:5 (Table IV.4.9). Perhaps with more weight on milk, these indexes would produce the desired changes in the shape of the curve without great loss in production.

It appears that the indexes of the third strategy could be feasible only if more weight were put on milk. These indexes in general do flatten the curve by decreasing b (the ascent), and by increasing <u>a</u> plus decreasing c.

IV.5. BLUP Solutions for Sires

The standard deviations for the BLUPs are in Table IV.5.1 for 305-day milk and the lactation curve parameters. The standard deviation of the BLUPs for 305-day milk was 267. The means of the BLUPs by definition, are zero. The range of the BLUPs for 305-day milk was 645 to -611, and 3.47 to -2.81 for peak yield. The top and bottom ranking sires for 305-day milk and peak yield were the same two sires.

Table IV.5.1 Standard deviations of BLUPs for 305-day milk yield and lactation curve parameters for 150 sires.

Variables or Parameters	Standard Deviation	
305-day milk	267.345	
a	.736637	
Ъ	.0107281	
с	1.59818 E-4	
b/c time of peak	4.82325	
Peak yield	.934166	
S ^A	62.4614	

A - S is $c^{-(b+1)}$, adjusted for a.

IV.5.1 Ranking sires by indexes

Using the weights of an index, an index for each sire was computed. These indexes were computed by a linear combination of the BLUPs for each trait and the weights of the index, i.e., $I_s = m_1 u_{1s} + m_2 u_{2s} + m_3 u_{3s}$. The subscript s refers to a specific sire, s = 1, ..., 150. This was done for a number of the indexes in the three strategies.

The rank of the 150 sires for several indexes was compared to their rank on milk alone. This comparison was done using Spearman's correlation of ranks. These correlations are in Table IV.5.2. The sires' rankings by indexes 1:1, 1:10 and 1:16 are not greatly different from those for milk alone. This is consistent with the change in milk expected when these indexes are used (Table IV.4.8). These three indexes represent large weighting on milk (1:1 and 1:16) or a heavy weight on peak (1:10) which is highly correlated to yield. In general, indexes of the first strategy had the highest correlations with rankings on milk alone. Indexes of the third strategy had the lowest, two of which were negative, 3:3 and 3:6. While those of the second strategy fell in the middle. This is consistent with the amount of genetic change in 305-day milk expected for the indexes when correlated responses, via the covariances, are considered [IV.4]. That is, for those indexes with genetic change in 305-day milk near that change expected when selecting for milk alone, the Spearman's correlation of ranks were high. Conversely, for indexes 3:3 and 3:6 the genetic change in milk was

	I	ndex				Correlation ^A
1:1	3m	1b	2peak	:		.9503
1:8	lm	6Ъ	6peak	:		.6435
1:10	lm	6 pea	k			.9023
1:13	lm	6Ъ				.2783
1:16	6m	1b				.9863
2:1	lm	15bc	lpeak	1S		.4379
2:5	lm	-10c	15bc	lpeak		.3078
2:6	lm	-10c	10bc	10peak		.7757
2:9	lm	-10c				.1963
3:3	lm	10a	-10b	-10peak	10S	2526
3:4	lm	10a	1Ъ	-10c	lpeak	.2323
3:6	lm	10a	-10b	-10c	-10peak	2178

Table IV.5.2 Rank correlations between sires ranked for milk only and other indexes.

A - Spearman's correlation of ranks.
considerably negative, -282 and -235 lbs and the ranked correlations were negative -.25 and -.22 for 3:3 and 3:6, respectively.

In order to consider the covariance between traits when computing I = m'u for each sire, the covariances must be incorporated in the BLUPs i.e., u. This can be done by expanding the random (sire) portion of the mixed model equations to include a variancecovariance matrix for each sire for the traits considered in the index. Multiple right hand sides i.e., one for each trait in the index are needed. This produces multiple BLUP solutions for each sire which are then combined by the weights, a to yield an index, I, for each sire. Computationally, this increases the random portion of the MME by a factor equal to the number of traits in the index. However, the method which was used in this study to combine the BLUPs for each trait into an index value, I, will yield the same value for I as the method just mentioned. The advantage in the procedure used in this study is the individual BLUPs can be computed ignoring the covariances and later combined into an index value, I.

V SUMMARY AND CONCLUSIONS

Currently, selection for milk production is based on total 305-day lactation yield. Although it is known that feed efficiency is the greatest and health costs are the highest in early lactation, these efficiency factors of a lactation have not been considered in selection. Considering the efficiency in early lactation, one may want to select cows that produce more in early lactation. On the other hand, if health costs are extensive during the high production, high stress period, then it may be economical to select cows which peak lower and later and are more persistent.

The purpose of this study is to fit first lactation records to Wood's equation and compute genetic estimates for the parameters <u>a</u>, b and c in the equation. Then, using selection indexes, change in the shape of the lactation curve along with 305-day milk yield will be selected jointly. This is an attempt to determine the flexability of the lactation curve shape and how it will affect total lactation yield.

Lactations of two year old cows in the Michigan DHI population were fit to the nonlinear form of Wood's equation resulting in parameters estimates for <u>a</u>, b and c for each cow. Using Shook's factors to adjust the first and/or second monthly tests, reduced the number of cows having curves with negative b values. This insured an ascent to the peak as opposed to lactations with estimated first day production greater than all subsequent test days. Also, two year olds are more likely to be increasing in production

at 305-days than later lactation cows. Therefore, using Shook's factors to compute production on the 305th day, based on the previous test, may underestimate it for some two year olds. It is likely, then, that adjusting the end of all two year old records will eliminate negative c values by causing a downward slope, but may do so in error. In the present study, this was done only for cows when their last test date was between 280 and 305 days.

Upon using Wood's equation, it is noted that b and c are not entirely independent. As c decreases, b increases and therefore peak yield and time of peak increase. It is also noted that as b increases, peak yield and yield after peak are greater.

Cows that increase in yield faster (larger b values) and maintain or decrease c (are more persistent) are expected to have a higher peak due to the relation in the computation for peak of the curve $(a(b/c)^{b}exp(-b))$. For two cows with the same <u>a</u> and b, peak yield dictates their persistency due to the relationship in the equation for peak. The cow with the higher peak will necessarily have a lower c and therefore, be more persistent. These conditions may not be true biologically. A more flexible equation would allow the ascent, peak, time of peak and persistency to be independent. This flexibility would improve the fit of lactation curves.

Computation of variance components using Best Linear Unbiased Prediction solutions from mixed model equations, and restricted maximum likelihood estimators in an iteration process was successful. Convergence occurred in ten or less iterations by using a relaxation step between iterations.

The heritability for milk production was less than is usually reported (.16). Heritabilities computed for <u>a</u> (.06), b (.09), c (.15), time of peak yield (.07), peak yield (.15) and S (.04) were all less than that for 305-day milk. Therefore, selection on milk yield alone produced greater genetic gain in 305-day milk yield than selecting for milk jointly with the lactation curve parameters.

Indexes including 305-day milk, the lactation curve parameters <u>a</u>, b and c, time of peak yield, peak yield and S, were set up for three strategies. The first strategy was to increase the amount of milk produced in the early part of lactation by increasing b and peak yield. The second strategy was an attempt to delay time of peak or decrease b, the ascent to the peak with or without considering persistency. The third strategy attempted to flatten the lactation curve by increasing <u>a</u>, decreasing peak yield and increasing persistency.

Indexes including milk, b and peak which are of the first strategy, resulted in nearly as much gain in 305-day milk as selecting for milk alone. These indexes have potential if it becomes desireable to increase yield in the peak part of lactation. In the first strategy, several restricted indexes were used to

restrict the genetic change in <u>a</u> and c so that b could be increased without decreasing <u>a</u> or increasing c. This was done in an attempt to increase peak yield $(a(b/c)^{b}exp(-b))$ by increasing b only. The progress made in 305-day milk by these restricted indexes was reduced considerably. Therefore, the restricted indexes were not useful.

Selecting for a delay in time of peak, the second strategy, in general, resulted in much lower gain in 305-day milk. When weights were 7, 3, 1, 1 for milk, time of peak, peak and S, respectively, the index, decreased the gain in milk somewhat less. However, the gain in time to peak was only .7 days per generation.

In the third strategy, selecting negatively for c, with equal weights for milk, greatly reduced the genetic change in 305-day milk (244 lbs) compared to selecting for milk alone (359 lbs). Conversely, selecting for milk alone had little influence on c. The correlation between 305-day milk and c suggest that high producing ability is not genetically related to persistency as measured by c. Selection for persistency is feasible, but if milk is to be maintained, it must have greater weighting than c.

Indexes which attempt to flatten the lactation curve, the third strategy, do so at the expense of 305-day milk, and with extreme weights, cause negative genetic gains in milk. These indexes selected positively for milk, S, and <u>a</u>, and negatively for b, c and peak. Therefore, the decrease in milk is to be expected. If these indexes are to be beneficial, the weights would have to be more in favor of milk. If flattening the curve results in decreasing stress and inputs substantially, then these indexes could be helpful. This is not likely, because increasing <u>a</u> and decreasing peak quickly decreases 305-day milk.

Of the three strategies, that first seems to be more in line with maintaining a reasonable gain in milk production while changing the shape of the curve. This is due to the positive relationship between peak yield and 305-day milk.

Indexes for each sire were computed by combining the BLUP estimates for each trait for each sire by the weights used in the selection indexes. This yielded an index for each sire. The sires were then ranked according to their indexes. Then, rank correlations for sires were computed between the ranking on each index and the ranking for milk alone. Ranking the sires using their BLUPs and the index weights suggest:

- (1) Rankings by indexes of the first strategy were very similar to rankings by milk alone, except when <u>a</u> and c were restricted. This suggests that most of the sires ranking high for milk alone also rank high for increasing peak yield and b.
- (2) In general, indexes which had genetic gain in milk close to that of selecting for milk alone, had high rank correlations.
- (3) Sires' rankings for indexes selecting to flatten the curve were poorly correlated to their ranks on milk alone and for some of these indexes negatively correlated.

The final step was to plot the shape of the lactation curve after 1, 5 and 10 generations of selection on each index. This

was done by putting the new genetic values for <u>a</u>, b and c, after selection by each index, into the equation $(y = at^{b}exp(-ct))$ and changing t from 1 to 305 days.

Two problems occur when attempting to plot the shape of the lactation curve after selection by indexes. The first is related to the nonlinear form of the equations used: $\hat{y}_{305-day\ milk}$ = $a_0 f^{305} t^b exp(-ct)$ dt. When <u>a</u>, b and c are changed linearly from generation to generation, the integral computed, i.e., estimated $305-day\ milk\ (\hat{y}_{305-day\ milk})$, changes nonlinearly. Therefore, to a small degree, the increments between generations are not equal for $\hat{y}_{305-day\ milk}$. This is due to the nonlinear relationship of the equation.

Second, the estimated values for 305-day milk computed by the integral of the new curves were not equal to the expected genetic change in 305-day milk when selecting on the indexes. These differences for some indexes were great. Both positive and negative differences occurred. One possible cause of this discrepancy is that the relationship between 305-day milk and some or all of the curve parameters is curvilinear. This means as genetic change in 305-day milk occurs in a linear fashion, the curve parameters change curvilinearly or vice-versa. Therefore, the correlations between 305-day milk and the parameters may change considerably when the selection process continues over 10 generations. Also, the relationship among some of the curve parameters may also be curvilinear. LIST OF REFERENCES

LIST OF REFERENCES

- Appleman, R. D., S. D. Musgrave, and R. D. Morrison. 1969. Extending incomplete lactation records of Holstein cows with varying levels of production. J. Dairy Sci. 52:360.
- Anderson, C. R. 1981. A biometrical and genetic study of Tribolium egg production curves as a model for lactation curves. Unpublished Ph.D. Thesis, University of Illinois.
- Bakker, J. J., R. W. Everett, and L. D. Van Vleck. 1980. Profitability index for sires. J. Dairy Sci. 63:1334.
- Barr, A. J., J. H. Goodnight, J. P. Sall, W. H. Blair and D. M. Chilko. 1979. <u>Statistical Analysis System (SAS) Users Guide</u>. SAS Institute Inc. Raleigh, N.C.
- Basant, S. and P. N. Bhat. 1978. Models of lactation curves for Hariana cattle. Indian J. Anim. Sci. 48:791.
- Brown, M. B., and A. B. Forsythe. 1974. Robust tests for the equality of variances. J. Amer. Stat. Assoc. 69:364.
- Cobby, J. M. and Y. L. P. Le Du. 1978. On fitting curves to lactation data. Anim. Prod. 26:127.
- Cochran, W. G. 1951. Improvement by means of selection. Proc. Second Berkely Sym. Math. Stat. and Prob. 449.
- Congleton, W. R., Jr. and R. W. Everett. 1980a. Error and bias in using the incomplete gamma function to describe lactation curves. J. Dairy Sci. 63:101.
- Congleton, W. R., Jr. and R. W. Everett. 1980b. Application of the incomplete gamma function to predict cumulative milk production. J. Dairy Sci. 63:109.
- Dickerson, G. E. 1969. Techniques for research in quantitative animal genetics. <u>Techniques and Procedures in Animal Science Research</u>. In Monograph of Amer. Soc. Anim. Sci. Champaign, ILL. pp. 36.

- Everett, R. W. 1975. Income over investment in semen. J. Dairy Sci. 58:1717.
- Everett, R. W., H. W. Carter, and J. D. Burke. 1968. Evaluation of the Dairy Herd Improvement Association record system. J. Dairy Sci. 51:153.
- Gaines, W. L. 1927. Persistence of lactation in dairy cows. Ill. Agric. Exp. Sta. Bul. 288. pp. 353.
- Gill, J. L. 1978. <u>Design and Analysis of Experiments in the Animal</u> <u>and Medical Sciences</u>. Vol. 1, Iowa State University Press, Ames, IA.
- Gooch, M. 1935. An analysis of the time change in milk production in individual lactations. J. Agr. Sci. 25:71.
- Gravert, H. O., and R. Baptist. 1976. Breeding for persistency of milk yield. Livest. Prod. Sci. 3:27.
- Grossman, M. 1970. Sampling variance of the correlation coefficients estimated from analyses of variance and covariance. Theoret. Applied Genetics. 40:357.
- Guest, P. G. 1961. <u>Numerical Methods of Curve Fitting</u>. Cambridge University Press, London, England. pp. 334.
- Hansen, L. B., R. W. Touchberry, C. W. Young, and K. P. Miller. 1979. Health care requirements of dairy cattle. II Nongenetic effects. J. Dairy Sci. 62:1932.
- Harris, D. L. 1963. The influences of error or parameter estimation upon index selection. <u>Statistical Genetics and Plant Breeding</u>. National Academy of Sciences National Research Council Publication 982. pp. 491.
- Harvey, W. R. 1972. General outline of computing procedures for six types of mixed models. Ohio State Univ. memeo.
- Hazel, L. N. 1943. The genetic basis for constructing selection indexes. Genetics. 28:426.
- Henderson, C. R. 1953. Estimation of variance and covariance components. Biometrics. 9:26.
- Henderson, C. R. 1963. Selection index and expected genetic advance. In <u>Statistical genetics and plant breeding</u>. NAS-NRC 982. pp. 141.

- Henderson, C. R. 1973. Sire evaluation and genetic trends. Proc. Anim. Breeding and Genetics Sympos. in Honor of Dr. Jay L. Lush, ASAS, Champaign, IL.
- Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics. 31:423.
- Johansson, I., and A. Hannson. 1940. Causes of variation in milk and butterfat yield of dairy cows. Kungl. Landbr. tidskr. 72 (6 1/2).
- Kellogg, D. W., S. Urquhart and A. J. Ortega. 1977. Estimating Holstein lactation curves with a gamma curve. J. Dairy Sci. 60:1308.
- Kempthorne, O. 1957. <u>An introduction to genetic statistics</u>. John Wiley and Sons, Inc., New York.
- Kempthorne, O. and A. W. Nordskog. 1959. Restricted selection indices. Biometrics. 15:10.
- Kendall, M. G., and W. R. Buckland. 1971. <u>A Dictionary of Statistical</u> <u>Terms</u>. Hafner Publishing Company, Inc. New York.
- Kendall, M. G. and A. Stuart. 1969. <u>Advanced Theory of Statistics</u>. Vol 1: Distribution theory. 3rd ed. Griffin Co. London.
- Keown, J. F., and L. D. Van Vleck. 1973. Extending lactation records in progress to 305-day equivalent. J. Dairy Sci. 56:1070.
- Lamb, R. C., and L. D. McGilliard. 1960. Variables affecting ratio factors for estimating 305-day production from part lactations. J. Dairy Sci. 43:519.
- Layard, M. W. J. 1973. Robust large-sample tests for homogeneity of variance. J. Amer. Stat. Assoc. 68:195.
- Lush, J. L. 1940. Intra-sire correlations or regressions of offspring on dam as a method of estimating heritability of characteristics. Thirty-third Annual Proceedings of the Amer. Soc. of Anim. Prod. 293.
- Madalena, F. E., M. L. Martinez and A. F. Freitas. 1979. Lactation curves of Holstein-Friesian and Holstein-Friesian x Gir Cows. Anim. Prod. 29:101.
- Madden, D. E., J. L. Lush, and L. D. McGilliard. 1955. Relations between parts of lactations and producing ability of Holstein cows. J. Dairy Sci. 38:1264.

- Mahadevan, P. 1951. The effects of environment and heredity on lactation. II Persistency of lactation. J. Agr. Sci. 41:89.
- Mao, I. L. 1971. The effect of parameter estimation errors on the efficiency of index selection and on the accuracy of genetic gain prediction. Unpublished Ph.D. Thesis. Cornell University.
- Mao, I. L. 1981. Lecture notes on multi-factor unbalanced data.
- Mao, I. L., J. W. Wilton and E. B. Burnside. 1974. Parity in age adjustment for milk and fat yield. J. Dairy Sci. 57:100.
- Marquardt, D. W. 1963. An algorithm for least-squares estimation of nonlinear parameters. J. Soc. Indust. Appl. Math. 1:431.
- McDaniel, B. T., R. H. Miller, E. L. Corley and R. D. Plowman. 1967. DHIA age adjustment factors for standardizing lactations to a mature basis. ARS-44-188. Dairy-Herd-Improvement Letter. Vol. 43, No 1.
- McGilliard, M. L. 1978. Net returns from using genetically superior sires. J. Dairy Sci. 61:250.
- Miller, R. H., and N. W. Hooven, Jr. 1969. Variation in part-lactation and whole-lactation feed efficiency of Holstein cows. J. Dairy Sci. 52:1025.
- Nelder, J. A. 1966. Inverse polynomials, a useful group of multifactor response functions. Biometrics. 22:128.
- Neter, J. and W. Wasserman. 1974. <u>Applied Linear Statistical Models</u>. Richard D. Irwin, Inc. Homewood, Illinois.
- Norman, H. D., P. D. Miller, B. T. McDaniel, F. N. Dickinson, and C. R. Henderson. 1974. USDA-DHIA factors for standardizing 305-day lactation records for age and month of calving. ARS-NE-40.
- O'Connor, L. K., S. Lipton. 1960. The effect of various sampling intervals on the estimation of lactation milk yield and composition. J. Dairy Res. 27:389.
- Pearson, R. E. 1976. Managing herd breeding: Economic considerations. Nat. Workshop on Genetic Improvement of Dairy Cattle. pp. 137.
- Schaeffer, L. R., C. E. Minder, I. McMillan and E. B. Burnside. 1977. Nonlinear techniques for predicting 305-day lactation production of Holsteins and Jerseys. J. Dairy Sci. 60:1636.

- Schneeberger, M. 1981. Inheritance of lactation curve in Swiss Brown Cattle. J. Dairy Sci. 64:475.
- Shanks, R. D. 1979. Relations among health cost, persistency, postpartum length, and milk production in Holstein cows. Unpublished Ph.D. Thesis, Iowa State University.
- Shanks, R. D., A. E. Freeman, P. J. Berger and D. H. Kelley. 1978. Effect of selection for milk production on reproductive and general health of the dairy cow. J. Dairy Sci. 61:1765.
- Shanks, R. D., P. J. Berger, A. E. Freeman, and F. N. Dickinson. 1980. Genetic aspects of lactation-curve parameters. Paper 126, presented at 75th Annual meeting of the American Dairy Science Association, Blacksburg, Virginia.
- Shanks, R. D., A. E. Freeman and F. N. Dickinson. 1981. Postpartum distribution of costs and disorders of health. J. Dairy Sci. 64:683.
- Shapiro, S. S. and M. B. Wilk. 1965. An analysis of variance test for normality (complete samples). Biometrika. 52:591.
- Shimizu, H. and S. Umrod. 1976. An application of the weighted regression procedure for constructing the lactation curve in dairy cattle. Japanese J. Zootech. Sci. 47:733.
- Shook, G. E. 1975. Outline of the proposed factors for improving accuracy of DHI estimates of lactation yield. Presented at National DHI Computing Workshop, Seattle, Washington, October.
- Smith, H. F. 1936. A discriminant function for plant selection. Ann. Eugenics. 7:240.
- Tallis, G. M. 1962. A selection index for optimum genotype. Biometrics. 18:120.
- Van Vleck, L. D. 1979. <u>Notes on the theory and application of</u> <u>selection principles for the genetic improvement of animals</u>. Cornell University. Third printing.
- Williams, J. S. 1962. The evaluation of a selection index. Biometrics. 18:375.
- Wood, P. D. P. 1967. Algebraic model of the lactation curve in cattle. Nature, Lond. 216:164.
- Wood, P. D. P. 1969. Factors affecting the shape of the lactation curve in cattle. Anim. Prod. 11:307.

- Wood, P. D. P. 1970. A note on the repeatability of parameters of the lactation curve in cattle. Anim. Prod. 12:535.
- Wood, P. D. P. 1972. A note on seasonal fluctuations in milk production. Anim. Prod. 15:89.
- Wood, P. D. P. 1976. Algebraic models of the lactation curves for milk, fat and protein production, with estimates of seasonal variation. Anim. Prod. 22:35.
- Yadav, M. C., B. G. Katpatal and S. N. Kaushik. 1977. Study of lactation curve in Hariana and its Friesian crosses. Indian J. Anim. Sci. 47:607.