ESTIMATION OF 305 DAY MILK YIELD FROM IRREGULAR MILK RECORDS

Thesis for the Degree of M.S.
MICHIGAN STATE UNIVERSITY
CURTIS C. MILLER
1966

THESIS

ESTIMATION OF 305 DAY MILK YIELD
FROM IRREGULAR MILK RECORDS

By

Curtis C. Miller

AN ABSTRACT

Submitted to the College of Agriculture
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Dairy

1955

Approved -

ABSTRACT

ESTIMATION OF 305 DAY MILK YIELD
FROM IRREGULAR MILK RECORDS

by Curtis C. Miller

Several methods of estimating milk yield for the lactation from various parts including several intervals between tests were evaluated on 58,840 completed records of Michigan Holstein cows calving between January, 1959, and January, 1961.

Three ratio procedures and linear regression were used to estimate total milk yield from the different parts for two monthly sets of bi-monthly data and three monthly sets of tri-monthly data.

The first ratio procedure consisted of averaged estimates where the estimates were those obtained by estimating total milk yield from the non-cumulative factors. These estimates provided error variances of 581,396, and 502 lb milk for the tri-monthly sets consisting of months 1-4-7-10, 2-5-8, and 3-6-9 respectively. The second ratio procedure consisted of weighting each non-cumulative estimate according to its variance and provided error variances for the three tri-monthly sets 1-4-7-10, 2-5-8, and 3-6-9 of 292,471 and 493 lb milk respectively. The third ratio procedure employed a single ratio of total to sum of production on groups of monthly test days to estimate total yield. This procedure

provided error variances of 279,459 and 475 lb milk for the tri-monthly sets consisting of months 1-4-7-10, 2-5-8, and 3-6-9 respectively. Bi-monthly estimates from the same ratio procedures provided error variances of 490,134 and 207 lb milk respectively for procedures 1,2, and 3 for the monthly set containing the first month of test and 415, 151, and 105 lb milk for the same three procedures for the monthly set consisting of months 2-4-6-8-10. All error variances were larger than those obtained by linear regression.

Regression factors were obtained for estimating total milk yield from various parts including several intervals between tests for four age-season groups

ESTIMATION OF 305 DAY MILK YIELD
FROM IRREGULAR MILK RECORDS

By

Curtis C. Miller

A THESIS

Submitted to the College of Agriculture
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Dairy

1966

Acknowledgements

I wish to express my sincere appreciation to Dr. Clinton B. Meadows for his encouragement and friendly guidance during the investigation of the problem and his helpful suggestions during the writing of the thesis.

The encouragement of Dr. Lon D. McGilliard was also appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (Continued ......)

LIST OF TABLES (Continued ....)

INTRODUCTION

Management and selection decisions for the dairy herd are based on production of each cow and the herd. Decisions on daily feeding are based on production at that particular time. Bulls are compared through the performance of their progeny where each animal's production is adjusted to a common age and constant length of lactation. In the culling of females from the herd, each animal is compared with the herd, again adjusting each animal to a common age and constant length of lactation. Traditionally, records have been adjusted to a common age using mature-equivalent factors and length of lactation has been 305 days, estimated from 10 monthly milk weights.

Considerable time could be saved in choosing between bulls and culling of the herd could be much more timely if an accurate estimate of the total lactation could be obtained from a part record. Less frequent milk weights and butterfat tests would reduce the cost of the testing program.

Total production has been estimated from part records by ratios and by regression. The ratio method relates the ratio of total production to a part for the lactation. The regression method takes into account some overall averageplus appropriate weights for each added part. To estimate total yield from cumulative and non-cumulative parts, uses either method. Regression coefficients have been formulated for estimating total yield from bi-monthly and tri-monthly tests. However, ratios have not been obtained for this purpose, nor have regression factors been computed for missing data.

Separate regression factors for various age-season groups are needed. There is also a need for computing appropriate factors for the various combinations ofsequential data and for tri- and bi-monthly sets with missing observations, where a monthly set consists of a particular group of months. There are three tri-monthly sets, one consisting of months 1-4-7-10, another consisting of months 2-5-8 and the third consisting of months 3-6-9. The bi-monthly sets contain months 1-3-5-7-9 and 2-4-6-8-10.

If ratios are to be used, there arises a question of handling additional data after the first estimate has been computed. Several possibilities exist for handling this situation; weight each non-cumulative estimate equally, weight the non-cumulative estimates in proportion to their relative precision, or form one ratio from all available data each time new data are available.

The objectives of this study are: (a) to compare the deviations from actual production of total production estimated from various parts by several methods, and (b) to seek ways to include longer intervals between tests without increasing errors of estimate.

REVIEW OF LITERATURE

Methods of estimating total lactation yield

Standard DHIA Method

The standard DHIA method of computing total lactation yield of a cow is to test the animal one time percalendar month within 3 days of a centering date for the herd. The amount produced on test day is applied back 15 days and ahead 15 days taking into account date of freshening and date dry or removed where appropriate. Production is cumulated to 305 days or to the end of the lactation. Campbell (1948), Erb et al. (1952), Flanagan (1965), McDowell (1927), O'Connor et al. (1960) and Rabild (1909) have reported accuracies of estimating total lactation yield from this method compared to the actual daily milk weights. The DHIA method is somewhat costly and laborious. Therefore, as early as 1915, workers were attempting to determine some method of estimating the total lactation yield of a cow from a partial record.

Yapp (1915) concluded that the standard DHIA method mentioned above, represented fairly accurately the producing abilities of cows and that a continuous seven day test was not a satisfactory criterion by which to judge a cow's total yield.

McDowell (1927) reported that the standard method varied on the average 2.91 percent from actual production. McCandlish and M'Vicar (1925) reported that the monthly method of testing produced results within two percent of the actual yield. Dick (1950) observed an error of 2.32 percent from actual production when cows were tested at 28 day intervals.

Using 96 bi-monthly, monthly, and six week intervals of recording production, Cunningham (1965) measured correlations between these estimates and actual 305-day milk production to be .99 for all three methods, .97 for bi-monthly and .98 for the other two methods. Actual data were milk weights each day for the entire lactation.

Bi-Monthly Tests

Erb *et al.* (1952), using 19 cows, reported that the calendar month method showed twice as much variation as the centering date (standard) method, but the former was not likely to be in error more than ± five percent for fat-corrected milk in 25 percent of the records, nor more than ± 12 percent in one percent of the records. The error could be reduced by less than one percent of that indicated for the 24 hour test by testing on a 48 hour basis and reduced less than two percent when the 96 hour test was compared with the 24 hour test.

The percent error in estimating milk yield exceeded by 25 percent of the records was 2.4, 3.4, 5.6, 7.4, and 8.8 for the 30, 60, 90, 120, and 150 day testing intervals, respectively. These data further substantiated the belief that cows tested at the first of a calendar month have an **advantage** over those always tested at the end of the month. These data on the 30 and 60 day intervals are in good agreement with McDowell (1927) and McCandlish and M'Vicar (1925).

McDowell (1927) found 3.8 percent error in estimates of fat when bi-monthly records were compared with monthly testing on 70 cows in the Minnesota Agriculture Experiment Station herd. This amounted to 6.65 lb fat per lactation on cows averaging 1.75 lb fat and five percent of the 70 cows had an error of 8.75 lb fat. As previously mentioned, McDowell (1927) concluded that the bi-monthly method would be in error 2.5 lb fat per lactation more than the standard method, but errors of both methods should be disregarded for all practical purposes.

Gifford (1930), reporting on more than 100 AR Holstein and Guernsey records, stated that 69 percent of the estimates were within one standard error of estimate in comparing bi-monthly with monthly records, 95 percent within two standard errors and 99 percent were within three standard errors of the estimate. For two combinations of bi-monthly testing reported mean deviations for fat of $-4.9 \pm .35$ for odd numbered months and $+2.4 \pm .38$ for even numbered months. For convenience he further divided the records by levels of fat production, using groupings of 300-400, 500-600, 700-800, and 900 lb. Correlation coefficients between monthly and bi-monthly records ranged from .956 to .997. From these data, Gifford concluded that a bi-monthly form of testing is satisfactory.

McKellip and Seath (1941) reported a correlation coefficient of .97 between monthly and bi-monthly records. They concluded that bi-monthly tests when used with daily weights were practically as accurate as monthly records made by centering the tests and not using daily weights. Copeland (1928) reported that records made under a bi-monthly method show little deviation from those made by the usual monthly methods. He used 500 Jersey records and found the deviation of all records to be 7.21 lb fat. He found 258 bi-monthly records that exceeded the monthly records and 242 that were lower. Alexander and Yapp (1949) used 684 cows of five breeds to compare bi-monthly with monthly tests and found that 43.27 percent of all cows were within two percent error of the actual, 29.54 percent of the total were within 2.1 to 5.0 percent of the actual, 20.03 percent within 5.1 to 10.0 percent of the actual, and 7.16 percent of the 684 cows were above 10.0 percent of the actual production. For all breeds, the average production was 11,894 lb milk with the average of the negative deviates being 11,802 and the positive 12,383 for the standard method. The bi-monthly method showed 11,453 lb milk for all cows, 12,431 lb for those above the average and 11,453 lb for those below average. These figures represented an error percentage from the standard method of -3.7 for the overall group, +0.4 for those above and -5.0 for those below the overall average. The authors concluded that bi-monthly testing is 92.84 percent as dependable as the standard method and this difference is not sufficiently large to exclude the practical use of bi-monthly testing.

The first data involving records from more than one herd were reported by Bayley _et_ _al_. (1952). They used 1,255 Holstein records of 305 days or less, but not less than 150 days, in 42 herds and reported a slightly greater percent error in the estimates than previous workers. Bayley compared the two sets of bi-monthly records with the standard DHIA method and found that the set which includes the first month on test is somewhat more accurate than the remaining set for both milk and butterfat. The relative reliability of the first set, where "relative reliability" was measured by the ratio of the mean squares of the two monthly sets, was reported as 101.0 percent and 104.0 percent for the first and second sets, respectively. The first set overestimated the milk record by an average of 69 lbs with a standard error of 2.8 lb. The second set underestimated the milk record by an average of 18 lb with a standard error of 3.4 lb. The frequency of errors larger than 10 percent for the first set was one in 78 and was one in 32 for the second set. The average percent error when bi-monthly records were compared with the standard method was 3.0 for all bi-monthly sets for milk and 4.0 for fat. Again, the first set was somewhat more accurate, being in error 2.8 percent for milk and 3.7 percent for fat while the second set was in error 3.2 percent and 4.4 percent for milk and fat. The workers concluded that bimonthly testing should be satisfactory for sire provings and population studies, but it may be unsatisfactory for individual lactation records.

Flanagan (1965) employed the centering date method to predict lactation records of six different lengths from seven, 14, 30, 42, 60, and 90 day intervals. From daily records of 367 lactations of s pring-calving Shorthorn cows and 147 lactations of spring-calving Friesians be found that different factors should be used to compensate for the bias introduced by the three months of calving, and for each of seven lactation length classes. For all lactation lengths combined, he reported standard errors of estimate as 9.35 imperial gallons for seven day intervals, 13.43, 22.48, 29.75, 40.32, and 69.60 imperial gallons for 14,30,42,60 and 90 day intervals, respectively for the Friesian breed. When he used the 305 day standard error as the base, the standard error of cows tested at 90 day intervals was 34.5 percent of the base.

Predicting Total Production from Parts

In the prediction of total lactation from one or more months of a cow's record the coefficient of correlation between the part and whole tells how valuable each part will be. The higher the correlation of production of a given month with the total yield, the more accurate is that segment in estimating the record.

Gaines (1927) found that a one or two day test during the fourth month of lactation provided the best indication of what the cow would produce during the lactation. Working with 80 Jersey and 80 Holstein records, Kennedy and Seath (1942) reported that production by first calf heifers during the first four months was a good index of what the first lactation would be and that it was useful to predict the relative production of the second lactation. Coefficients of correlations between production in single months and total production for first lactation for Holsteins were .62 for the first month and .78 for the fourth month. The correlation between complete first and second lactations was .54.

Madden et al. (1955) reported repeatabilities, heritabilities, and genetic correlations for monthly and cumulative periods of milk production. Repeatabilities for milk and fat production in single months and for cumulative milk and fat production were 0.41, 0.32, 0.57, and 0.51, respectively. By intra-sire regression of daughters on dam, heritabilities were .076 for 275-305 to .390 for 1-30 days production for monthly data and .344 for 1-120 days to .632 for 1-274 days of cumulative production. Genetic correlations between cumulative parts and total milk and fat production were larger than 0.90. Selecting on the cumulative part record would improve production nearly as much as selecting on the complete record itself, with efficiency values ranging from 0.74 to unity.

Reece (1942), working in New Jersey with 70 cows in the Experiment Station herd, concluded that the average fat test at the end of the second month of lactation was a good measure of the ability of a cow to secrete milk fat. Furthermore, he found that almost as much accuracy was available by the end of the sixth month of lactation as for 10 full months. He reported correlations with total production of 0.53 for the first month, .74, .80, .85, .87, and .93 for two through six cumulative months of the lactation.

Voelker (1957) extended 1636 two-year-old records from the college herd at South Dakota and reported correlations of total production of .68 for the first month to .99 for nine cumulative months on test. Correlation coefficients increased to .89 for five cumulative months then remained the same for the next four months. The cumulative milk yield for the first 70 days of lactation was a good indication of total lactation yield, correlation coefficient of .80 reported by Rendel (1957), using 3109 production records from six breeds in Great Britain. Madden _et_ _al_. (1959) found the correlations between non-cumulative test-day production and total yield were higher for younger cows and that the highest values occurred between the fourth and seventh months, ranging from .71 to .93. Correlations between production on test dates varied from .38 to .89 for cows under two years old on 2X milking and from .23 to .88 for cows three years old and older. The highest correlations occurred between adjacent months. These results were similar to those reported earlier by Kennedy (1942), Voelker (1957), and Reece (1942) in the United States and Rendel (1957) in Great Britain. Lamb (1962) also reported relatively high correlations compared with the other months for the fourth, fifth, and sixth months individually with the total lactation yield.

The younger cows tended to be more persistent than the older animals; correlations of first month with total milk declined from .68 for the first lactation, to .62 for the second, and to .54 for the third lactation and over. The middle months varied less than the first and last months of the lactation, but the younger cows were always highest and the older animals always lowest in production. This was also true for butterfat values. VanVleck and Henderson (1961d), in ascertaining the efficiency of intra-herd regression factors compared to those derived by ignoring herd effects, reported similar findings. They found correlations of .67 between first month and milk yield for total lactation for 9036 Holstein cows, and the correlation increased to .90 for the fifth month of lactation, then decreased to .52 for the tenth month. Again, the fourth and sixth months were also higher than all others except the fifth month, when herd effects were ignored. The correlation between part and whole for cumulative milk production increased from .67 for the first month to .99 for the ninth month. On an intra-herd basis, non-cumulative data provided correlations of .53 for the tenth month and .57 for the first month. Cumulatively, the correlation coefficients were .57 for the first month and .99 for the ninth month. These values were similar to those later reported by Lamb (1962). However Lamb's data were recorded by lactation number and the older animals had lower values, particularly in the first five months, than did the first lactation animals. The same pattern for fat was evident in Lamb's (1962) data. Miller (1963) used Lamb's non-cumulative factors for milk and fat and reported a correlation of .92 between actual herd averages from 553 herds and averages estimated from single days of test of 19,000 cows at various stages of the lactation. The correlation was .82 for milk and .62 for fat production when computed for individual cows.

Extension Factors and Their Derivations

Two methods have been used to derive factors for estimating production from parts for the total lactation (305 days). Both methods provide factors for estimating total production from data in a single month or from production cumulated from more than one month, using the ratio of total production to a particular part. The projection equation is $y = cx$ where *y* is the estimated 305-day record, *c* is the ratio of the completed lactation to the part, and *x* is the production for the part. The regression equation can be represented by $y = a + bx$ where *y* is the estimated 305-day record, *a* is the point where the regression line intercepts the *y* axis, and *b* is the regression coefficient measuring the average change in y for each unit change in *x*, and *x* is the actual production for the part.

Ratios

Turner and Ragsdale (1924) formulated ratios to estimate 305-day fat production from AR records of lesser length. Cannon *et al*. (1924) selected at random 15 cows of three breeds and computed the errors resulting from estimating a 305-day record. They observed an average difference of 8.8 lb fat with the estimated records being larger for the "College Herd" factors and a negative value of 3.0 lb fat when factors derived from DHIA data were used to extend records. They concluded that the errors in records calculated by ratio factors are small, and the factors could be used successfully.

Turner and Ragsdale used over 3,000 Holstein lactation AR records to ascertain appropriate age and length factors which would allow them to compare daughters of 229 Holstein sires. Cannon used 400 Holstein records from the Iowa State College herd and 1289 lactation records of five breeds from the Iowa Dairy Herd Improvement Association to calculate cumulative factors to extend partial records to a 305 day basis. The conversion factors were based on the average rate of decline by using the ratio of total production during 365 days to the production at various monthly intervals.

Turner (1926) illustrated how the exponential law could be employed in expressing in a quantitative form the persistency of milk and fat secretion and its application to the analysis of experimental data. His argument was that from a given maximum production for a month, the cause of variation in total yield of milk is due to the variation in the rateof decrease. He suggested that an index of persistency could be obtained from the ratio between total yearly yield and the maximum month's production. When this ratio equals twelve for a yearly test or ten for ten months, the persistency percent is 100. This ratio applied to the maximum month's production provided an estimate of the 305-day production when ten months were used in the total. Turner concluded that when all other conditions are uniform, the monthly milk or fat production during the lactation period after the maximum is passed is a constant percentage of the preceeding month's production.

Madden et al. (1959) considered age, frequency of milking, and production levels in formulating factors to estimate total lactation records from parts. They used the first ten tests from lactations which had production recorded for at least ten different test dates and divided the data into lactations with the same milking frequencies (2X or 3X). They found that lactation curves for 2X and 3X for the same lactation appeared parallel, with the 3X curve being higher than the 2X curve. Earlier Madden et al. (1955) found lactation curves parallel for cows calving at less than three years of age, and for three, four, five, and six year olds, but the older cows differed from the younger animals in achieving a higher maximum but declining more rapidly. The curves crossed at about nine months in production. Average production on first test day for monthly age groups through six years of age differed significantly, but the means for ages under three years and means for three years and older did not differ within their respective groups. There was overlapping of average production between groups only at the ages of 35 to 37 months. Of those animals calving between 35 and 36 months of age the percent of first lactations initiated was approximately equal to the percent of total lactations initiated. Twelve percent of the second lactations were initiated prior to 36 months and ten percent of the first lactations after 35 months. For this reason, they classified the ages into less than 36 months and equal to or greater than 36 months.

Madden _et al_. (1959) concluded that the ratio method may appeal intuitively to many and that it is easy to develop and use although it tends to underestimate total production of low-producing cows and overestimate total production of high producing cows since the ratio method corrects only for incompleteness of the lactation and does not take into account the incomplete repeatability of the parts of the lactation.

Lamb (1959) completed a comprehensive study of variables affecting the relationship of the total to part production. He used 16,272 complete lactation records of four breeds in Michigan to study age and lactation number, season of calving, herd effect, and breed effect. First calf heifers did not decline in production as rapidly as older cows during the last months of the lactation. By grouping cows according to first, second, and third or more lactations and according to age by less than 36 months, 36 to 48 months, 48 to 60 months, and 60 months and over, he found the variability in ratios of total to part was larger for lactation number than for age. This was true particularly for the first and last months of lactation. The component of variance for age during the center months was larger than the component for lactation; thus, age would be a better indicator for these months. He further stated that age correction factors should be used to extend first lactation records initiated after 36 months of age, since the factors for first lactation would overestimate production. Age correction factors also would be more realistic for animals with second lactations initiated prior to 36 months, since the factors based on lactation would tend to underestimate production. In all other cases, factors based on either age or lactation should work equally well.

Lamb (1962) substantiated the previous work and suggested that at least two sets of age correction factors be used. Separate factors could be used for records initiated at less than 36 months, 37 to 47 months, and over 48 months, but it appears reasonable to use one set of factors for all cows calving at ages over 36 months. If one set of ratio factors were used, records for older cows would be overestimated while those for younger cows would be underestimated. These results were similar to those reported by Madden _et_ _al_. (1959) for Holstein HIR data.

Lamb (1959) classified records by four seasons, October-December, January-March, April-June, and July-September. Season appeared to account for almost as much variation as age, and Holsteins were least influenced by season of calving.

Two seasonal groups were used to develop ratio factors - November-April and May-October. Lamb's later work did not substantiate earlier results; he found the season adjustment was less important than adjustment for age. He also rearranged the two seasonal groups to include April-July and August-March since differences between the factors of these groups were larger than between factors of the two previous groups.

Lamb's (1959) work indicated no sizable variation in the relationship of total to part production among various classes of records. However, Jerseys tended to show the most herd differences of the four breeds included.

Aulerich (1965) reported ratio factors for predicting total production from terminal incomplete lactations differed from those for non-terminal records. The terminal incomplete lactations were initiated at a lower milk yield than those of cows completing their records in the same age-season group, and they declined at a faster rate. A separate set of ratio factors was developed for cows having voluntarily terminal incomplete records. Cows removed voluntarily were culled for low production, old age, dairy purposes, or hard milking.

The DHIA Newsletter (1965) includes a set of projection factors for estimating 305-day milk production. These factors were based on a total of 162,191 DHIA lactations combining Michigan and Iowa data. The authors recommend the use of two age factors, disregarding season effects, since previous research indicates that age is the most important cause of variation within breed in the projection factors.

VanVleck and Henderson (1961b) classified cows into 60 age classes, three seasons of calving and ten stages of lactation to ascertain the differences beteen ratio factors for each of the above effects. Their analysis of 177,575 Holstein records indicated that ratio factors for estimating total lactation yield from cumulative monthly records must be constructed simultaneously taking into account age at calving, season of calving, month of production, and whether the record is milk or fat. Ratio factors for adjusting monthly records to a common age and season and for estimating a total lactation from a single monthly record must be constructed in the same manner. This method would require 3600 factors, too many to be practical. Therefore, the authors suggest using six-month age intervals which would reduce the number of factors to 1200. They remarked that correcting for age but not for season and then analyzing records within seasons would not be adequate because definite differences exist among seasons.

Regression

Gowen and Gowen (1922) to provide some means for making the seven day test more meaningful to farmers, used regression factors to estimate total lactation yield from the seven day totals. They reported correlation coefficients of -.115 to .835 (with a weighted mean of .598) between seven-day and 365-day totals for various age groups. In addition to formulating prediction equations for the same lactation, they predicted the succeeding lactation from a given record. They found that a seven day test predicted the 365-day record of which it is a part more accurately than it predicts subsequent lactations.

Fritz <u>et</u> <u>al</u>. (1960) investigated the importance of breed, herd, lactation number, season of calving, and age at calving on the relationship between milk or fat production on test day and corresponding production for the complete lactation. They computed appropriate regression factors for estimating 305-day records from parts. In this study, they used 11,420 Michigan DHIA-IBM records, including only those animals having at least ten consecutive test days. They reported that variation in fat production among seasons was significant only for the first test day of production and that herd differences in milk production were significant for the first test day of production. An examination of age and lactation numbers revealed that only seven percent of the first lactations were initiated after 36 months of age and five percent of the second lactations prior to 36 months of age.

Intra-herd factors were compared visually with inter-herd factors and, although these results were not tested for significance, the authors reported a marked similarity between the two sets of factors, especially for milk production. Since variation due to herd differences was not significant except for the first month of production, the authors suggested that it may not be necessary to derive extension factors on an intra-herd basis to achieve sufficient maximum accuracy in extending records to 305 days. Correlations between cumulative test day production and the 305-day total, ignoring herd effects, also were reported and compared closely with others previously mentioned.

To ascertain if separate extension factors are needed for low or high producing cows , Madden _et_ _al_. (1959) fit quadratic regression equations of total on cumulative part production for various age-milking frequency groups. They reported that multiple correlation values were within .006 of the correlations reported when production levels were not considered. They developed regression factors for extending cumulative part records at two milking frequencies and two age groups. Except for young cows in the first month of the lactation, the differences between frequencies of milking within ages were non-significant.

VanVleck and Henderson (1961a) used 9,036 Holstein records from 375 New York herds and reported regression factors for extending (1) single monthly records, (2) sequential test day data, and (3) cumulative test day records to a ten month total on an intra-herd basis. Correlations between monthly and total production ranged from .53 for the tenth month to .85 for the fourth, fifth, and sixth months when used singly. Correlation coefficients for cumulative production ranged from .57 for the first month to .99 for the ninth month, and these were in general agreement with others previously reported. The authors reported regression factors for extending part lactation records to a ten-month total for those situations when data for the first month are missing.

In a later paper, VanVleck and Henderson (1961d) compared the efficiency of using intra-herd regression factors to those where herd effects are ignored. The records were adjusted for age at calving and season of calving. The relative efficiency (E) was defined as 100 (V_1/V_2) where V_1 was the residual variance of the sum of the first ten test records not accounted for by total regression and V_2 was the residual variance of total yield not accounted for by intra-herd regression. The relative efficiencies of the intra-herd factors ranged from 102 percent for the first nine sequential months to 158 percent for the tenth month alone. For most cases, ignoring herd effects was 10-20 percent less efficient than considering herd effects. Residual variances for regression ignoring herd effects for single month's records ranged from 1,736 for the fifth month to 6,446 for the tenth month. For sequential months, the variance ranged from 51 for the first nine sequential months to 4,585 for the first month. For cumulative months the variance was highest for the first month and lowest at the ninth month. Bi-monthly residual variances, ignoring herd effects, were 166 and 171 for sequential monthly sets of odd numbered months and even numbered months of lactation respectively. For three tri-monthly sets variances were 679, 428, and 393. Cumulative functions of these data gave the following variances: 214 and 213 for sets one and two of bi-monthly records and 904, 457, and 421 for the three tri-monthly sets. The authors concluded that prediction of ten month milk yield by regression ignoring herd effects is more practical for most situations because of simplicity although the accuracy is slightly less in all situations than for the intra-herd predictions.

Comparison of the Ratio and Regression Methods

Madden *et al.* (1959) stated that the choice between the ratio and regression methods depends on the purpose for which the method is to be used and the ease of use. They indicated that the ratio method may have more appeal to many and would be easier to develop and use. The variation in the total production estimated by ratio is more nearly like the variation in actual total production. In addition to the adjustment for incompleteness accomplished by the ratio method, the regression method also adjusts for the unidentified sources of variation which make the part larger or smaller than average, and the total estimated by regression varies less than the actual total. The differences between these two methods are largest during the early months of the lactation.

Lamb (1959) reported that the value of the regression procedure is in its ability to correct for the incomplete repeatability of various portions of the lactation, whereas the advantages of the ratio procedure are its direct measure of the relationship of the partial lactation to its total, its simplicity, ease of calculation, and its availability from a single lactation.

Harvey (1956) found that separate equations are needed for each stage of lactation when the regression procedure is used because of the changing parameter values. He noted that the estimated intercept approached zero as the length of lactation increased, causing similarities to exist between the ratios and regression coefficients. The differences between the two methods can be expressed as $(b-c)(x-\bar{x})$ where b is the linear regression of total production on a part, c is the ratio of the total to part, x is the actual partial milk production, and $\bar{x}$ is the mean of the partial milk production for the appropriate breed-age-season group. Madden et al. (1959) found the differences (b-c) to be negative and largest in magnitude during the early (one-three) cumulative months, negative but smaller in magnitude during the middle months (four-six), and positive and small during the longer (seven-nine) cumulative months.

SOURCE OF DATA

Approximately 58,840 completed records of Holstein cows were selected from more than 500,000 records tested in Michigan DHIA between January, 1959, and October, 1961. The analyzed records in this study were obtained by selecting completed Holstein records which were included in a previous study (Aulerich, 1963). In the previous work, records of milk production were obtained from approximately 2,500,000 monthly reports of cows tested in Michigan DHIA from January, 1959 to October, 1961. Only lactations initiated after January 1, 1959 and identified by herd number, cow number, date of calving, age at calving and breed were included. Each test day milk weight was recorded in tenths of pounds, and records in which the first test occurred more than 50 days after calving were excluded. From this group, records were chosen to include only those cows that had consecutive monthly production from calving until going dry. Records of cows leaving the herd for various reasons were excluded from this analysis. Likewise, lactations with data missing for one or more months were excluded.

METHODS AND RESULTS

Record Classification

Data were analyzed as a single group for the primary purpose of this study to compare various methods of estimating production in the total lactation. For computing regression coefficients, the data were categorized by two age groups (cows calving at less than 36 months of age and those calving at 36 months or over) and two seasons of calving (cows calving from April through July and those calving from August through March). The basis of these classifications is primarily Lamb's (1959 and 1962)work. All ratio procedures were limited to include all cows in milk on test date and the comparison between regression and the ratio procedures was conducted on this basis rather than for each of the various age-season groups.

Table 1 shows the classification of data as used for computing the age-season regression factors.

Measuring Relationships of Parts to the Whole

The value of any part in estimating the whole is determined by how closely the part is related to the whole. Simple product-moment correlations were used to measure this relationship. When several parts were included in the relationship, multiple correlation coefficients were used to express the relationship between the parts combined and the whole,

TABLE 1

Distribution of Records by Age and Season

Age at Calving	Percent
$<$ 36 months	36.5
$\geq$ 36 months	63.5
Season of Calving	
April - July	23.4
August - March	76.6
Age-Season of Calving	
$<$ 36 months - April - July	8.6
$<$ 36 months - August - March	27.9
$\geq$ 36 months - April - July	14.8
$\geq$ 36 months - August - March	48.7

partial correlation coefficients were used to express the highest order partial correlation coefficient between the parts (X_i) and the whole (Y).

Most data used in previous reports have had the restriction that the first ten consecutive months of test be included. To determine whether the relationships of part to whole were similar in these cases to those instances where animals terminate their record by going dry prior to the completion of the first ten test dates, correlations were computed for both situations.

Aulerich (1965) suggested that ratios may differ for terminated records after plotting the lactation curves for the two groups. She stated that involuntary terminal records were similar to the complete records and that these data should be extended using the non-terminal factors. She further reported that voluntarily terminal incomplete lactations should be extended using another set of factors. Therefore, she developed a separate setof ratios for these records. This could also be accomplished for animals going dry prior to 305 days.

Table 2 **shows** the percent and numbers of animals in milk for each of the ten test dates.

These data indicate that 18 percent of the cows going dry prior to 305 days terminate their records between the ninth and tenth month of lactation.

Table 3 shows the simple correlations between milk produced on individual test days and the production for the total lactation for records including the first ten tests.

Table 4 provides corresponding correlations between various parts and the total production for all cows in milk on test day.

Estimating 305-Day Records from Various Parts of the Lactation

Ratios

The ratios, which were extensively developed by Lamb (1962), Madden *et al.* (1955) and Aulerich (1965), are a popular form of extending records. Two forms of ratios have been developed, (1) cumulative ratio factors which extend the total production to date to the whole, and (2) non-cumulative ratios which extend single month's production to the total.

A linear function of the ratios to extend records from two or more nonadjacent test dates can be expressed as $R_i = (R_j^{-1} + R_k^{-1})^{-1}$ where R_i = ratio to extend the sum of amounts produced on two non-adjacent months.

R_j = ratio used to extend on non-cumulative basis the jth test

R_k = ratio used to extend on non-cumulative basis the kth test.

This ratio may be useful where some tests are missing.

Another method of using ratios is to average estimates from individual test days. This could be accomplished by dividing the estimate by the number of ratios used, thus, weighting equally each ratio. This can be expressed as

TABLE 2

Percent Cows in Milk on Each of 10 Test Dates

	Month of Test						
	4	5	6	7	8	9	10
Percent	99.8	99.5	99.3	99.0	98.3	97.4	82.0
	Percent Cows Reported Dry the First Time for Each of 10 Test Dates						
Percent	1.34	1.28	1.01	1.76	3.81	20.2	70.4
Total	17.99						

TABLE 3

Correlations Between Milk Produced on Single Test-Days for Records with 10 Tests

Month of Test	Month of Test										
	1	2	3	4	5	6	7	8	9	10	Total
1	1.0										.75
2	.82	1.0									.85
3	.75	.87	1.0								.88
4	.69	. 82	.88	1.0							.90
5	.64	.76	.82	.87	1.0						.91
6	.58	.60	.76	.81	.87	1.0					.90
7	.50	.61	.67	.74	.80	.86	1.0				.87
8	.39	.49	.56	.62	.69	.76	.84	1.0			.81
9	.24	.32	.33	.44	.50	.58	.66	.80	1.0		.67
10	.13	.20	.24	.29	.35	.42	.50	.63	.81	1.0	.53

TABLE 4

Correlations Between Milk Produced on Single Test-Days (Dry Test Days Excluded)

Month of Test	Month of Test										
	1	2	3	4	5	6	7	8	9	10	Total
1	1.0										.71
2	.81	1.0									.82
3	.75	.87	1.0								.85
4	.69	.82	.88	1.0							.88
5	.64	.76	.82	.87	1.0						.89
6	.56	.68	.75	.81	.87	1.0					.89
7	.47	.59	.65	.72	.78	.86	1.0				.87
8	.34	.45	.51	.58	.65	.73	.83	1.0			.80
9	.20	.29	.35	.41	.47	.55	.65	.81	1.0		.66
10	.13	.20	.24	.29	.35	.42	.50	.63	.81		.53

$(R_1X_1+R_2X_2 + R_3X_3) / 3$ where R_1, R_2, R_3 are the non-cumulative ratio factors for any three test days and X_1, X_2, X_3 are the corresponding amounts of milk for each of the test days.

To weight each estimate of total production by the inverse of its error variance can be expressed as follows:

$\Sigma \lbrack R_i X/(1-r_{iy}^2) \hat{\sigma}_{iy}^2 \rbrack / \Sigma \lbrack 1 / 1-r_{iy}^2) \hat{\sigma}_{iy}^2 \rbrack$ where

R_i = non-cumulative factor for extending records for the ith test date

r_{iy}= correlation between the ith part and the total

$\hat{\sigma}_{iy}^2$= total variance of 305-day lactations.

Standard DHIA Method

The DHIA Handbook provides a centering date table for bi-monthly testing which functions in the same manner as the regular monthly testing program. Each bi-monthly testing period is divided as nearly as possible into two equal groups of time, half of the time occurring prior to the test date and half the time falling after the test date. Many states have used a bi-monthly system to provide a low cost testing program. The model for the DHIA procedure can be written: $\hat{y} = B_1X_1 + B_2X_2 + B_2X_3 + B_2X_4 + B_3X_5$ where the values of B_1 and B_3 depend on the stage of lactation when the animal is first tested and the X_i represent the milk weights on test days. The value of B_2 is 2.0 since each test represents two monthly tests.

Regression Coefficients for Estimating Total Yield

The least squares normal equations were solved to estimate the desired regression coefficients. The right and left hand sides of these equations were made up of the total sums of squares and cross-products corrected for means.

Algebraically, the estimates are obtained from: $b_i = \sum_{j=1}^{c} c_{ij} \cdot M_{jy}$ where b_i is the estimated partial regression coefficient of total production on production in part i of the lactation.

c_{ij} is the element of the inverse of the matrix of sums of squares and products involving production parts i and j.

M_{jy} is the corrected sum of products of total production and production in part j.

The standard error of the regression coefficient is estimated as $S_{b_i} = \sqrt{c_{ii} S^2}$ where S^2 is the error variance or $(1-R_I^2)\hat{\sigma}_y^2$ where R_I^2 is the square of the multiple correlation and $\hat{\sigma}_y^2$ is the estimated variance of the total yield.

VanVleck and Henderson (1961d) concluded that this method is more practical for most situations if herd effects are ignored.

An example of the use of regression factors to estimate total yield from a test-day of a single record is as follows: $\hat{Y}$ = total lactations for the breed + b_i (test day production - breed average production for that test day or $\hat{Y} = \bar{Y} + b_i(X_i = \bar{X}_i$).

Regression coefficients and their standard errors, standard errors of estimated yields, and correlations of various parts with total production were computed for single months, sequential months, cumulative months, and different types of both bi-monthly and tri-monthly testing.

Most work to date has been based on records containing at least ten test dates, with the exception of that of Madden et al. (1955) who included completed records of 243 or more days in length. To ascertain whether or not this type of restriction affects the factors, the regression factors in this study were computed in the two ways, one including those records containing at least ten consecutive test dates and the other also including records terminated by dry period prior to 305 days.

Table 5 shows the means and standard deviation of monthly test day records and the cumulative production for those animals in milk at least ten consecutive test dates. A total of 48,257 records was used in this category, with the average age being 47 months. Table 6 contains corresponding values for data including all cows in milk on test date.

Tables 7 and 8 contain regression coefficients for estimating total milk yield for the lactation from cumulative test data for both situations. The standard errors of estimates are smaller for the data where ten consecutive test dates are used than for records of all cows in milk on test day.

Tables 9 and 10 compare similar regression coefficients when the production in single months is used to project the total yield. Since all cows in milk on the tenth test day are the same animals in milk ten consecutive test dates, the results for these two cases are identical, but results for the two methods diverge from the tenth to the first test month's data.

Table 11 contains the regression coefficients for extending records when the first month's data are missing. The number of animals included in the estimates is the same as the number of animals in milk at least ten consecutive dates.

Thus far, all factors have been simple regressions where only one part is used to estimate total production. Tables 12 and 13 are results of multiple linear regressions where two or more parts are involved. This can be represented as follows: $Y = \quad + b_1X_1 + b_2X_2 \ldots b_kX_k$ where represents $\bar{y}-b_1\bar{X}_1-b_2\bar{X}_2- \ldots b_k\bar{X}_k$

Values are computed for both sets of data, records available for the first ten consecutive test dates and all data available on a particular test date. Table 12 lists data for situations where records for one or more sequential test dates at the end of lactation are missing. Table 13 provides corresponding values for all cows in milk on test date.

Table 14 contains values used to extend part records on a sequential basis when data are missing in the first part of lactation.

TABLE 5

Means and Standard Deviations of Monthly Test Day Records of Milk Produced on Single and Cumulative Test Days (10 Test Days)

Month of Test	1	2	3	4	5	6	7	8	9	10
Single Months	51.3	51.0	46.9	43.3	40.4	37.9	35.3	32.4	28.1	23.7
Std. Dev.	12.6	12.7	11.9	10.9	10.1	9.5	9.0	8.7	9.0	9.2
Cum.Mos. Means	51.3	102.3	149.2	192.5	232.9	270.8	306.1	338.5	366.6	390.4

Total	390.4
Std.Dev.	84.0
305 Day (Actual)	11907

TABLE 6

Milk Produced on Single Test Days (Dry Test Days Excluded)

Month of Test	1	2	3	4	5	6	7	8	9	10
Mean	51.2	50.6	46.4	42.7	39.8	37.1	34.3	30.9	26.9	23.7
Std.Dev.	12.7	12.8	12.0	11.0	10.2	9.7	9.4	9.4	9.5	9.2

TABLE 7

Regression Factors for Estimating Total Milk Yield from Cumulative Test-Day Records (10 Test Days)

Month of Test	1	2	3	4	5	6	7	8	9
b	4.97	2.91	2.12	1.72	1.48	1.31	1.19	1.11	1.05
Std.b	.020	.009	.005	.004	.003	.002	.001	.001	.001
r	.75	.84	.88	.91	.93	.95	.97	.98	.99
$\hat{\sigma}_e$	55.9	46.2	39.9	34.8	30.0	25.4	20.6	15.2	8.3

TABLE 8

Regression Factors for Estimating Total Milk Yield From Cumulative Test Day Records (Dry Test Days Excluded)

Month of Test	1	2	3	4	5	6	7	8	9
b	4.98	2.93	2.15	1.74	1.49	1.32	1.21	1.12	1.06
Std.b	.020	.009	.006	.004	.003	.002	.002	.001	.001
r	.71	.80	.85	.88	.91	.93	.95	.97	.99
$\hat{\sigma}_e$	62.9	53.6	47.6	41.7	36.6	31.6	26.1	19.3	10.8

b = Regression coefficient
r = Correlation between that month and the total
$\hat{\sigma}_e$ = $\sqrt{\text{Resid Var.}}$
Std b = Std. error of regression coefficient

TABLE 9

Regression Factors for Estimating Total Milk Yield From A Single Monthly Test Using Data Only From Cows in Milk at Least Ten Test Dates

Month of Test	1	2	3	4	5	6	7	8	9	10
b	4.97	5.58	6.23	6.96	7.54	7.94	8.10	7.79	6.27	4.80
Std.b	.020	.016	.015	.015	.016	.017	.021	.026	.032	.035
r	.75	.85	.88	.90	.91	.90	.87	.81	.67	.53
$\hat{\sigma}_e$	55.9	44.7	39.8	36.5	35.2	36.6	41.2	49.2	62.4	71.5

TABLE 10

Regression Factors for Estimating Total Milk Yield From A Single Monthly Test Record Using Data From All Cows In Milk on Test Date

Month of Test	1	2	3	4	5	6	7	8	9	10
b	4.98	5.68	6.34	7.09	7.66	8.03	8.05	7.34	5.94	4.80
Std.b	.020	.017	.016	.016	.016	.017	.019	.023	.028	.035
r	.71	.82	.85	.88	.89	.89	.87	.80	.66	.53
$\hat{\sigma}_e$	62.9	51.5	46.5	41.8	39.3	39.2	42.7	51.4	63.5	71.5

TABLE 11

Regression Factors for Estimating Total Milk Yield From Cumulative Test Day Records When First Month Records are Missing

Month of Test	10	9	8	7	6	5	4	3	2
b	4.80	3.05	2.50	2.15	1.87	1.64	1.44	1.26	1.11
Std.b	.035	.017	.011	.007	.005	.003	.002	.001	.001
r	.53	.63	.73	.81	.88	.92	.96	.98	.99
	71.5	65.4	57.3	48.8	40.6	32.7	24.8	17.1	9.4

TABLE 12

Regression Factors for Estimating Total Milk Yield
From Sequential Test-Dates Using Data Only From
Cows In Milk At Least Ten Test Dates

Sequential Months of Test	1	2	3	4	5	6	7	8	9	R	$\hat{\sigma}_e$
1 b	4.97									.75	55.9
Std. b	.020										
2 b	1.09	4.70								.85	44.0
Std. b	.028	.027									
3 b	.67	1.70	4.11							.90	37.1
Std. b	.023	.031	.029								
3 b	.65	.99	1.65	3.91						.93	31.3
Std. b	.020	.027	.030	.028							
5 b	.68	.85	1.02	1.68	3.63					.95	26.0
Std. b	.016	.022	.025	.028	.025						
6 b	.77	.86	.84	1.11	1.53	3.32				.97	21.1
Std. b	.013	.018	.021	.023	.024	.021					
7 b	.85	.93	.85	.94	1.04	1.42	2.91			.98	16.5
Std. b	.010	.014	.016	.018	.019	.020	.017				
8 b	.92	.99	.93	.93	.94	1.00	1.11	2.59		.99	11.2
Std. b	.007	.010	.011	.012	.013	.013	.014	.011			
9 b	.98	1.00	1.00	.97	.98	.98	.99	1.04	1.85	1.00	5.4
Std. b	.003	.005	.005	.006	.006	.006	.007	.007	.005		

TABLE 13

Regression Factors for Estimating Total Milk Yield From Sequential Test-Dates When Last Months Data Are Missing Using Records of All Cows in Milk on Test Date

Sequential Months of Test	1	2	3	4	5	6	7	8	9	R	δ_e
1 b	4.98									.71	62.9
Std.b	.020										
2 b	.90	6.02								.75	58.7
Std.b	.034	.033									
3 b	.42	1.78	4.35							.87	44.5
Std.b	.025	.034	.032								
4 b	.43	.99	1.71	4.17						.90	37.9
Std.b	.022	.030	.033	.030							
5 b	.49	.85	1.03	1.68	3.92					.93	32.1
Std.b	.018	.025	.029	.031	.028						
6 b	.63	.88	.85	1.03	1.46	3.72				.95	26.7
Std.b	.015	.021	.024	.026	.028	.024					
7 b	.78	.95	.87	.86	.92	1.29	3.44			.97	20.8
Std.b	.012	.016	.019	.020	.022	.023	.018				
8 b	.90	1.02	.95	.90	.88	.91	1.08	2.92		.99	13.9
Std.b	.008	.011	.012	.014	.015	.015	.015	.011			
9 b	.95	1.03	1.04	.96	.94	.95	.93	1.21	1.86	1.00	7.8
Std.b	.005	.006	.007	.008	.008	.009	.008	.008	.006	.	

Bi-Monthly and Tri-Monthly Results

It appears from previously mentioned results that it may be necessary to compute regression coefficients to extend records which have only various portions of the data normally available for computation of the tri-monthly and bi-monthly records. For example, if a cow has no production on her first test day, she still fits into the first bi-monthly or tri-monthly set but with the first test date's data missing.

If a full complement of monthly data are needed to extend the production to a 10 month basis, records with missing monthly results would be excluded from the program. However, if regression factors are available for extending a particular incomplete tri-monthly set, the data could be salvaged. The number of combinations of months varies from those situations where all months are available to those where only one is available. For tri-monthly testing, the monthly set using four test dates (1-4-7-10) would have a total of 15 combinations of incomplete data when single months are included; the other two sets have a total of seven combinations each. However, since the single months and the complete compliment of dates are already computed, it is necessary only to solve for ten additional factors for the 1-4-7-10 combination and three additional factors for each of the other two sets, 2-5-8 and 3-6-9 combinations.

Table 15 lists the simultaneous values for both tri-monthly and bi-monthly testing systems. These data contain all records available on test date. The corresponding cumulative figures for both bi- and tri-monthly testing are shown in Table 16. Table 17 shows the regression values to be assigned for the various "incomplete" monthly sets for each of the three sets of tri-monthly data. The number of combinations that could result for each of the two bi-monthly sets is quite large (at least 30). However, many of these will fall into the single month category or into one of the above listed tri-monthly results.

Average Production Within Age and Season

Means of total production were highest for animals initiating lactations at 36 or more months of age and from August to March, and were lowest for those animals freshening during April-July at ages less than 36 months. The younger cows reached peak production at lower levels and did not diminish in production as rapidly as the older animals. Table 18 gives the means and standard deviations for the four age-season groups for ten months during the lactation.

Table 19 shows the regression factors for estimating total yield from cumulative test days data. Differences in these factors point to differences in production levels of aged cows and young animals. Larger factors are required for the younger animals early in the lactation; but as production accumulates, factors for different ages converge.

Table 20 is a table of regression factors for the four age-season groups for extending cumulative part records to a ten month basis when data are missing for the first month.

TABLE 14

Regression Factors for Estimating Total Milk Yield From Sequential Test-Dates When First Month's Data Are Missing Using Records From All Cows in Milk on Test Date

Sequential Months of Test		2	3	4	5	6	7	8	9	10	R	$\hat{\sigma}$
1	b									4.80	.52	71.5
Std.	b									.035		
2	b								6.61	-.41	.67	62.3
Std.	b								.053	.052		
3	b							7.33	.63	-.09	.81	49.1
Std.	b							.043	.054	.041		
4	b						6.05	1.99	.40	.33	.88	39.1
Std.	b						.036	.047	.044	.033		
5	b					5.00	2.07	1.04	.51	.61	.93	31.0
Std.	b					.030	.037	.037	.035	.026		
6	b				4.01	1.98	1.19	.84	.67	.80	.96	24.0
Std.	b				.022	.028	.029	.029	.029	.020		
7	b			3.11	1.76	1.26	.95	.85	.79	.91	.98	17.9
Std.	b			.016	.020	.022	.022	.022	.020	.015		
8	b		2.34	1.42	1.19	1.02	.94	.92	.93	.94	.99	12.4
Std.	b		.010	.013	.014	.015	.015	.015	.014	.010		
9	b	1.67	1.15	1.04	1.04	.99	.97	.99	.98	.96	.99	7.2
Std.	b	.005	.007	.008	.008	.009	.009	.009	.008	.006		

TABLE 15

Regression Factors for Estimating Total Milk Yield
from Sequential Bi-Monthly and Tri-Monthly Test-Day Records

Month of Test		1	2	3	4	5	6	7	8	9	10	R	$\hat{\sigma}_e$
						Bi-Monthly							
Set 1	b	1.37		2.06		1.96		2.04		2.42		.99	12.5
Std.	b	.006		.009		.011		.011		.007			
Set 2	b		2.23		2.04		2.04		1.99		1.41	.99	11.7
Std.	b		.007		.011		.012		.011		.008		
						Tri-Monthly							
1	b	1.81			3.14			3.11			1.88	.98	15.0
Std.	b	.008			.011			.012			.009		
2	b		2.62			2.84			3.78			.97	19.2
Std.	b		.010			.014			.011				
3	b			3.59			3.16			2.64		.97	19.1
Std.	b			.010			.015			.010			

TABLE 16

Regression Factors for Estimating Total Milk Yield From Cumulative Bi-Monthly and Tri-Monthly Test-Day Records

	Bi-Monthly		Tri-Monthly		
Monthly Sets	1	2	1	2	3
b	1.93	2.00	2.52	3.00	3.19
Std. b	.001	.001	.002	.003	.003
R	.99	.99	.98	.97	.97
$\hat{\sigma}_e$	14.4	13.0	17.2	20.3	20.0

TABLE 17

Regression Factors for Estimating Total Milk Yield From Sequential Tri-Monthly Test Day Records When Data From One or More Test Dates are Missing

Month of Test	1	2	3	4	5	6	7	8	9	10	R	$\hat{\sigma}_e$
Sequence												
1-4-7 b	1.58			2.89			4.64				.96	24.2
Std. b	.001			.016			.015					
1-4-10 b	2.02			4.89						2.98	.99	20.9
Std. b	.010			.012						.010		
4-7-10 b				4.60			3.17			1.66	.96	22.1
Std. b				.014			.018			.013		
1-7-10 b	2.97						5.15			1.75	.96	24.3
Std. b	.010						.016			.014		
1-4 b	1.36			6.00							.89	40.0
Std. b	.018			.021								
1-7 b	2.71						6.34				.94	30.1
Std. b	.010						.013					
1-10 b	4.60									3.98	.86	42.6
Std. b	.016									.021		
4-7 b				4.24			4.50				.95	28.2
Std. b				.015			.018					
4-10 b				6.31						2.63	.94	28.2
Std. b				.012						.015		
7-10 b							7.53			1.12	.88	40.2
Std. b							.023			.023		
2-5 b		2.31			5.47						.92	34.2
Std. b		.017			.021							
2-8 b		3.94						4.96			.96	25.0
Std. b		.009						.012				
5-8 b					5.49			3.52			.94	29.0
Std. b					.016			.017				
3-6 b			3.16			5.11					.94	30.2
Std. b			.016			.020						
3-9 b			5.21						3.70		.95	25.9
Std. b			.010						.012			
6-9 b						2.38			7.49		.96	34.8
Std. b						.019			.016			

Table 21 shows both correlations between part and total and the regression factors for extending single month's data to a ten month basis. Correlations are higher for younger animals than for older animals within a particular season of calving. The differences are larger in the early months of lactation than for the middle or last months of the record.

Table 22 contains regression factors for estimating total lactation yield from sequential test-day data for the four age-season groups. Again the correlations between the estimate and actual are higher during the early parts of lactation for younger animals than for older cows.

Table 23 lists the regression factors for estimating total milk yield from sequential bi-monthly data for four age-season groups. Correlations between these estimates and the actual values are also included.

Table 24 contains regression coefficients for estimating total yield from tri-monthly data. The correlations between the estimate and actual production using the first, fourth, seventh, and tenth month's data are as high or higher than the other two test sets for all four age-season groups, while the third set is as low or lower than the other two for the same age-season group. Regression factors to estimate total production from both bi-monthly and tri-monthly cumulative data are included in Tables 25 and 26.

TABLE 18

Test Date Means and Standard Deviations for Four Age-Season Groups

(a)	≥ 36 months			August-March							
Months	1	2	3	4	5	6	7	8	9	10	Total
Single Month	54.9	54.4	50.0	46.2	43.3	40.4	37.6	33.9	28.6	23.2	412.5
Std. Dev.	11.4	11.8	11.5	10.6	10.1	9.7	9.4	9.3	9.6	9.7	83.4
Cumulative Month Means	54.9	109.3	159.3	205.5	248.8	289.2	326.8	360.7	389.3	412.5	
(b)	**< 36 months**			**August-March**							
Months	1	2	3	4	5	6	7	8	9	10	Total
Single Month	42.8	43.2	40.5	38.4	36.7	35.1	33.4	31.5	28.7	24.9	355.2
Std. Dev.	10.4	10.6	10.1	9.4	8.9	8.5	8.1	7.9	8.2	8.5	75.5
Cumulative Months Mean	42.8	86.0	126.5	164.9	201.6	236.7	270.1	301.6	330.3	355.2	
(c)	**≥ 36 months**			**April-July**							
Months	1	2	3	4	5	6	7	8	9	10	Total
Single Months	57.3	56.7	50.8	45.2	40.8	37.1	34.0	30.9	26.9	23.1	402.8
Std. Dev.	11.7	12.0	11.1	10.4	9.5	9.0	8.5	8.3	8.6	9.2	78.2
Cumulative Months Means	57.3	114.0	164.8	210.0	250.8	287.9	321.9	352.8	379.7	402.8	

TABLE 18 Continued ...

(d)	<36 months			April-July							
Months	1	2	3	4	5	6	7	8	9	10	Total
Single Months	45.8	45.6	41.5	37.8	35.0	33.1	31.4	29.8	27.4	25.0	352.4
Std. Dev.	11.1	11.5	10.4	9.7	8.8	8.2	7.6	7.5	7.7	8.3	74.5
Cumulative Months Mean	45.8	91.4	132.9	170.7	205.7	238.8	270.2	300.0	327.4	352.4	

TABLE 19

Regression Factors For Estimating Total Milk Yield From Cumulative Test-Day Records for Four Age-Season Groups

	Age-Season Groups			
Cumulated Months of Test	≥ 36 months Aug.-Mar.	< 36 months Aug.-Mar.	≥ 36 months April-July	< 36 months April-July
1	5.20	5.56	4.55	4.97
2	3.14	3.21	2.87	2.91
3	2.28	2.32	2.14	2.13
4	1.84	1.85	1.74	1.73
5	1.56	1.56	1.50	1.49
6	1.37	1.37	1.34	1.33
7	1.23	1.23	1.23	1.22
8	1.13	1.13	1.14	1.14
9	1.06	1.06	1.06	1.06

TABLE 20

Regression Factors for Estimating Total Milk Yield From Cumulative Test Day Records When First Month Records are Missing for Four Age-Season Groups

Month of Test	≥ 36.months Aug.-Mar.	<36 months Aug.-Mar.	≥ 36 months April-July	< 36 months April-July
2	1.10	1.11	1.10	1.11
3	1.23	1.25	1.23	1.26
4	1.39	1.43	1.39	1.43
5	1.59	1.65	1.60	1.63
6	1.82	1.91	1.78	1.85
7	2.10	2.24	2.06	2.18
8	2.45	2.65	2.46	2.62
9	3.05	3.30	3.14	3.40
10	4.97	5.35	5.09	5.42

TABLE 21

Regression Factors for Estimating Total Milk Yield from a Single Monthly Test Record for Four Age-Season Groups

Month of Test		≥36 months Aug.-Mar.	<36 months Aug.-Mar.	≥36 months April-July	<36 months April-July
1	b	5.20	5.56	4.55	4.97
	r	.71	.77	.68	.74
2	b	5.83	6.13	5.26	5.45
	r	.83	.86	.81	.84
3	b	6.30	6.63	5.93	6.23
	r	.87	.89	.84	.87
4	b	6.93	7.27	6.49	6.82
	r	.89	.91	.86	.89
5	b	7.39	7.72	7.27	7.65
	r	.90	.91	.88	.90
6	b	7.75	8.06	7.66	8.19
	r	.90	.91	.83	.90
7	b	7.80	8.24	7.90	8.50
	r	.88	.89	.86	.87
8	b	7.42	7.92	7.76	8.24
	r	.83	.83	.82	.83
9	b	6.08	6.58	6.58	7.24
	r	.70	.71	.73	.75
10	b	4.97	5.35	5.09	5.42
	r	.58	.60	.60	.61

TABLE 22

Regression Factors for Estimating Total Milk Yield from Sequential Test Day Data for Four Age-Season Groups

Age-Season Group			(a)	≥36 months		August-March				
Seq. months of test	1	2	3	4	5	6	7	8	9	R
1	5.20									.71
2	1.27	4.94								.84
3	.89	1.93	4.01							.89
4	.81	1.21	1.56	3.91						.92
5	.81	1.00	.93	1.74	3.50					.95
6	.83	.98	.82	1.14	1.50	3.26				.96
7	.90	.98	.83	.95	1.05	1.38	2.90			.98
8	.96	1.00	.92	.93	.94	1.01	1.11	2.57		.99
9	.99	1.01	.99	.99	.97	.98	1.01	1.02	1.33	.99

			(b)	<36 months		August-March				
	1	2	3	4	5	6	7	8	9	R
1	5.56									.77
2	1.27	5.11								.87
3	.85	1.84	4.30							.91
4	.80	1.15	1.74	3.90						.94
5	.80	.97	1.16	1.63	3.57					.96
6	.87	.90	.87	1.10	1.67	3.19				.97
7	.95	.92	.84	.88	1.21	1.53	2.75			.98
8	.97	1.03	.89	.85	1.03	1.09	1.21	2.43		.99
9	.97	1.02	1.00	.94	1.00	1.00	1.00	1.03	1.79	.99

TABLE 22 Continued ...

Age-Season Group			(c)	≥ 36 months		April-July				
Seq.months of test	1	2	3	4	5	6	7	8	9	R
1	4.55									.68
2	1.42	4.23								.82
3	1.02	1.78	3.67							.87
4	.94	1.04	1.73	3.46						.91
5	.85	.76	1.16	1.50	3.54					.94
6	.95	.77	.85	1.20	1.42	3.42				.96
7	.98	.87	.87	1.04	.98	1.30	3.15			.98
8	.95	.95	.97	1.02	.96	.90	1.03	2.62		.99
9	.99	.99	.99	.96	1.02	.99	.96	1.01	1.95	.99
			(d)	< 36 months		April-July				
	1	2	3	4	5	6	7	8	9	R
1	4.97									.74
2	1.29	4.46								.85
3	.74	1.97	3.77							.89
4	.71	1.07	1.80	3.59						.92
5	.73	.69	1.17	1.79	3.62					.95
6	.80	.67	.92	1.25	1.50	3.59				.97
7	.86	.77	.90	1.07	1.11	1.65	2.95			.99
8	.92	.92	.92	1.03	.98	1.05	1.24	2.62		.99
9	1.00	.93	.97	.97	.99	.97	.99	1.13	1.89	.99

TABLE 23

Regression Factors for Estimating Total Milk Yield from Sequential Bi-Monthly Test Day Records For Four Age-Season Groups

Age-Season Group	≥36 months Aug.-Mar.		<36 months Aug.-Mar.		≥36 months April-July		<36 months April-July	
Seq. month of test	Monthly sets*							
	1	2	1	2	1	2	1	2
1	1.40		1.35		1.38		1.42	
2		2.16		2.11		2.22		2.21
3	1.99		1.93		2.08		1.97	
4		2.03		1.96		2.08		1.96
5	2.02		2.06		2.03		2.00	
6		2.08		1.95		2.03		2.02
7	2.03		2.03		2.04		2.04	
8		1.96		2.22		1.90		2.08
9	2.30		2.42		2.25		2.52	
10		1.45		1.37		1.48		1.34
R	.99	.99	.99	.99	.99	.99	.99	.99

* 1 is 1-3-5-7-9 monthly set
2 is 2-4-6-8-10 monthly set

TABLE 24

Regression Factors for Estimating Total Milk Yield from Sequential Tri-Monthly Test Day Records For Four Age-Season Groups

Age-Season Group	≥36 months Aug.-Mar.			<36 months Aug.-Mar.			≥36 months April-July.			<36 months April-July		
Seq. month of testn	Monthly Sets*											
	1	2	3	1	2	3	1	2	3	1	2	3
1	1.76			1.77			1.80			1.85		
2		2.56			2.43			2.72			2.60	
3			3.63			3.28			3.62			3.51
4	3.18			2.97			3.19			3.00		
5		2.96			2.75			2.99			2.64	
6			3.25			3.14			3.16			3.17
7	3.08			3.35			3.02			3.27		
8		3.53			4.03			3.46			4.05	
9			2.60			2.85			2.57			2.73
10	1.91			1.77			1.91			1.80		
R	.98	.93	.95	.98	.97	.97	.99	.93	.98	.98	.93	.97

* 1 is 1-4-7-10 monthly set
2 is 2-5-8 monthly set
3 is 3-6-9 monthly set

TABLE 25

Regression Factors for Estimating Total Milk
Yield from Cumulative Bi-Monthly Test Day Records

Age-Season Group	≥36 months Aug.-Mar.		<36 months Aug.-Mar.		≥36 months April-July		<36 months Apri.-July	
Month of Test	Monthly Set							
	1	2	1	2	1	2	1	2
b	1.93	1.97	1.93	1.95	1.93	1.98	1.94	1.96
R	.99	.99	.99	.99	.98	.98	.99	.99

TABLE 26

Regression Factors for Estimating Total Milk
Yield from Cumulative Tri-Monthly Test Day Records

Age-Season Group	≥36 months Aug.-Mar.			<36 months Aug.-Mar.			≥36 months April-July			<36 months April-July		
Month of Test	1	2	3	1	2	3	1	2	3	1	2	3
b	2.53	2.97	3.11	2.49	2.94	3.11	2.51	3.01	3.18	2.50	2.97	3.19
R	.98	.97	.97	.97	.97	.97	.98	.98	.98	.98	.97	.97

COMPARISON OF METHODS

Madden *et al.* (1959) have shown that differences between regression and ratio factors are largest during the early cumulative months of lactation and smallest in the latter stages (7-9 cumulative months). To express these differences as a relationship between the estimated and actual performance, the standard error of estimate was used for regression for ratios or other schemes, the variance of the difference, Harvey (1956) explained that the increased variance of the difference between estimated and actual total production becomes $(b-c)^2 \sigma X^2$ where X^2 equals the variance of cumulative part production.

Correlations between estimated and actual production and standard errors of the estimates are shown in Table 27 for four methods, regression, average estimates by ratios, proportionally weighted ratios, and the cumulative non-adjacent month ratios. Correlations and standard errors of estimates also are included when total production is estimated by the DHIA method as explained earlier in this paper. Only those sets having a complete complement of test dates were used for the DHIA procedures.

Ratio and Regression Methods

Non-cumulative and cumulative ratio and regression factors were compared to point out where the two varied and by how much. The regression factors have ignored herd effects; therefore, these results will not coincide with

TABLE 27

Comparisons of Deviations of Differences Between Estimates and Actual Values for Various Methods Used to Estimate Ten Month Total Milk Production

	Monthly Sets of Data																			
Method	1-4-7-10		1-4-7		1-4		1		4		7		1-7		1-10		1-7-10		1-4-10	
Regr (a)	15.0	.98	24.2	.96	40.0	.89	62.9	.71	41.8	.88	42.7	.87	30.1	.94	42.6	.86	24.3	.96	20.9	.99
Ratio (b)	16.7	.98	29.9	.94	47.1	.86	71.2	.71	46.5	.88	51.3	.87	36.5	.91	45.3	.86	29.0	.94	26.5	.95
Ratio (c)	17.1	.98	24.7	.96	42.0	.89	----	---	----	---	----	---	37.6	.91	55.9	.81	32.6	.93	24.0	.96
Ratio (d)	24.7	.96	26.8	.95	44.8	.87	----	---	----	---	----	---	32.6	.93	59.1	.79	40.2	.89	34.2	.92
DHIA	16.7	.98																		

	4-10		4-7		7-10		4-7-10		10	
Regr	28.2	.94	28.2	.95	40.2	.88	22.1	.96	71.5	.53
Ratio	39.2	.91	29.1	.95	62.1	.81	31.3	.94	131.8	.53
Ratio	36.5	.92	29.2	.95	55.3	.84	28.5	.95	----	---
Ratio	42.3	.90	30.5	.94	72.5	.75	43.0	.85	----	---

TABLE 27 Continued

	Monthly Sets of Data													
Method	2-5-8		2-5		2-8		5-8		2		5		8	
Regr	19.2	.97	34.2	.92	25.0	.96	29.0	.94	51.5	.82	39.3	.89	51.4	.80
Ratio	21.4	.97	36.3	.91	28.0	.95	32.8	.94	56.9	.82	44.5	.89	70.0	.80
Ratio	21.7	.97	36.1	.92	28.7	.95	32.7	.94	----	---	----	---	----	---
Ratio	19.9	.97	37.3	.92	27.2	.95	35.7	.93	----	---	----	---	----	---
DHIA	22.2	.97												
	3-6-9		3-6		3-9		6-9		3		6		9	
Regr	19.1	.97	30.2	.94	25.9	.95	34.8	.96	46.5	.85	39.2	.89	63.5	.66
Ratio	21.8	.97	34.0	.93	28.8	.95	47.4	.88	57.4	.82	45.3	.89	102.1	.66
Ratio	22.2	.97	31.6	.94	34.6	.93	43.0	.90						
Ratio	22.4	.97	32.4	.94	36.4	.92	52.9	.86						
DHIA	21.4	.97												

(a) Regr. is the partial regr. coefficient.
(b) Ratios is the ratio formed from cumulative non adjacent test months data
(c) Non-cumulative ratio according to its error variance
(d) Non-cumulative ratio estimates averaged

those of Madden *et al.* (1959) concerning the difference (b-c).

For the non-cumulative and cumulative factors, (b-c) is always negative. However, this value is larger for the first three months of cumulative production than for the last three months where c and b converge.

The value of (b-c) is negative for all months, largest the last three months, and most similar the third, fourth and fifth months of the lactation. Tables 28 and 29 show these results.

Distribution of Errors

Any method that estimates the total from part should provide not only high correlation of estimated and actual production but also, if the absolute values are to be used, estimates that are close to the actual values. The standard errors of estimates provide some picture of this since they indicate the frequencies of various magnitudes of deviations of estimates from actual production. The standard errors for the regression method are less than those for any other method, and the standard errors are smaller for the fifth and sixth single months than for any other single months. The standard errors for regression factors used in projecting cumulative data are lower when only those animals are included that have at least ten consecutive test dates than when all cows in milk on test date are included in the results. The pattern is similar for sequential data. The results for cumulative data where only cows in milk ten months were included compare favorably to the data of Fritz et al. (1960) where inter-herd regression factors were used, with the standard errors in the present data being slightly lower for all months. Standard errors also were smaller than those reported by VanVleck and Henderson (1961d).

Another comparison of the various methods used to extend records to a 305-day basis may be obtained from the distribution of deviations of estimates from actual values. For situations where the deviations do not adhere to a normal distribution, this method will provide more information concerning the precision of the various methods than will the standard errors alone.

Table 30 is a compilation of the percent of the deviations falling in each of eight categories for three methods used in extending records to a 305-day basis. The Chi Square Test for goodness of fit was used to test whether these data fit the normal distribution. For all three metnods the lack of fit was significant at the 0.01 probability level; i.e., these deviations do not follow a normal distribution. The tests were made using averages of the three tri-monthly sets rather than testing each tri-monthly set separately, since the probability of rejecting each hypothesis when it was true for all three monthly sets would then be greater than the chosen level of 0.01.

The data show that the distributions of deviations for the three monthly sets obtained by the regression method

TABLE 28

A Comparison of Non-Cumulative Regression and Ratio Factors

	Months of Test									
Method	1	2	3	4	5	6	7	8	9	10
Regr. Factors	4.98	5.68	6.34	7.09	7.66	8.03	8.05	7.34	5.94	4.80
Ratios	7.66	7.68	8.28	8.92	9.53	10.23	11.08	12.34	14.33	16.91

TABLE 29

A Comparison of Cumulative Regression and Ratio Factors

	Months of Test								
Method	1	2	3	4	5	6	7	8	9
Regr. Factors	4.98	2.93	2.15	1.74	1.49	1.32	1.21	1.12	1.06
Ratios	7.66	3.84	2.62	2.03	1.67	1.43	1.27	1.15	1.07

TABLE 30

Frequencies of Estimated Value Deviations from Actual Values for Various
Methods of Projecting Tri-Monthly Testing Data to a Ten Month Total

Method	Normal Distribution Categories								
	>3 Std. Dev.	2 to 3 Std. Dev.	1 to 2 Std. Dev.	0 to 1 Std. Dev.	0 to -1 Std. Dev.	-1 to -2 Std. Dev.	-2 to -3 Std. Dev.	Less than -3 Std. Dev.	Std. Dev.
					33 (a) Monthly set 1-4-7-10				
Ratio	.3%	2.3%	14.2%	38.4%	31.3%	10.7%	2.3%	.5%	16.7
Wtd. Ratios	.4%	2.6	14.5	36.7	31.1	11.5	2.4	.7	17.1
Regression	.03%	.24	13.9	37.9	34.2	13.0	.5	.16	15.0
					33 (b) Monthly set 2-5-8				
Ratio	.2	1.7	11.4	38.7	35.5	10.5	1.7	.3	21.8
Wtd. Ratios	.3	1.7	11.0	38.9	38.0	8.7	1.2	.2	22.2
Regression	.2	1.3	10.5	38.1	37.0	10.9	1.7	.3	19.1
					33 (c) Monthly set 3-6-9				
Ratio	.2	1.6	11.1	37.9	38.5	9.4	1.1	.2	21.4
Wtd. Ratios	.5	2.7	14.6	36.0	31.7	11.6	2.4	.5	21.7
Regression	.2	1.4	11.4	39.3	34.3	10.9	2.0	.6	19.2
					33 (d) Average				
Ratio	.23	1.87	12.23	38.33	35.10	10.20	1.70	.33	
Wtd. Ratios	.30	2.33	13.36	37.20	33.60	10.63	2.00	.46	
Regressions	.14	.96	11.92	38.44	35.18	11.60	1.41	.31	

were more alike than for the other two methods. For example, the mean ± one standard deviation for the regression method included 72.1 percent, 75.1 percent, and 73.6 percent of the observations for the three sets whereas the ratio method included 69.7 percent, 74.2 percent, and 76.4 percent. The weighted ratios included 68.8 percent, 76.9 percent and 67.8 percent for the three monthly sets, being about as inconsistent as the ratios. The ratio method overestimates the production 46.8 percent of the time, weighted ratios 46.7 percent and regression 43.5 percent of the time.

DISCUSSION

Relationship Between Various Parts and Total Yield

The primary interest is in tri-monthly testing. Therefore, while data are included for sequential, cumulative, single months, bi-monthly, and tri-monthly data, most discussion will center around the tri-monthly figures. Results are similar for the other methods and will be discussed only when they differ from the tri-monthly findings.

Correlations of .89 between single monthly parts and total are highest for the fifth and sixth single months of production when all records available on test day are used. The corresponding values for data restricted to only animals in milk ten test dates are .91 and .90 for the fifth and sixth months, respectively. In the latter case, the fourth month is as highly correlated to the total as the sixth month. All correlations of part with whole are higher for the restricted data than for data where all animals are included. This would indicate that the record of an animal going dry prior to having ten full months of production tends to deviate both plus and minus from the average lactation curve more than the records of counterparts. Regression and correlation coefficients reported here for animals in milk at least ten test dates are similar to those reported previously by Madden et al. (1959) but the correlations are somewhat higher than those reported by VanVleck and Henderson (1961d). Largest differences between these figures and VanVleck's occur early in the lactation, with values reported here somewhat larger for the early months (e.g. .75 vs. .67 for the first month).

Lamb (1962) reported correlations between monthly production and total lactation yield in three different age groups. The highest correlation for any age group between the first month and total yield was .68 for the animals in their first lactations. Table 4 shows this value for all cows (average age 47 months) to be .71. Similar differences occurred in the other nine single months. Lamb (1962) found that the highest correlations occurred in the fourth and fifth months of production but these were slightly less than values reported in this study (.86 vs. .88 for the fourth month and .86 vs. .89 for the fifth month.) Lamb (1962) included only animals completing the first ten months of production.

As could be expected, the highest correlations between parts occurred between adjacent months of the lactation and the lowest occurred between the parts most distant. The lowest correlation involving the first month is with the tenth month (.13). These correlations of parts are of value in determining what combination of months will provide the best information about the total. Perhaps this can best be seen in one of the tri-monthly combinations where the monthly set is 3-6-9. The standard error of estimate when all three months were used was 19.1; however, the standard errors for the 3-9, 6-9, and 3-6 combinations were 25.9, 34.8, and 30.2 lb, respectively. Of the three months used, th e sixth month is most closely correlated with the total .89, .85, and .66 for months 6, 3 and 9, respectively. However, when the sixth month is dropped from the combination, the standard error is smaller than when either the third or ninth month is dropped. The squared multiple correlation (R^2) is affected similarly and is .84 when month three is excluded, .89 when month nine is excluded and .91 with month six excluded. These differences may be explained partially by production in the sixth month being more highly correlated with production in the third or ninth month than is production in the third with the test in the ninth.

Correlations between various parts are generally higher for the "ten consecutive test" groups than for the group including all animals in milk on test day. This is especially true for the sixth through ninth months of lactation, and in particular when correlated with the earlier months in the lactation.

Precision of Various Methods Used to Extend Records

Ratios

Ratios for cumulative test days in non-adjacent months which are linear combinations of the ratios for individual test days, were used to extend various parts to a ten-month basis. Correlations between these estimates and actual production were similar to but never greater than those obtained by multiple linear regression. Variance of the deviations of estimates from actual production was used as a basis of comparison. Weighting the individual estimate with the inverse of its "variance of estimate" also was used to combine ratios to extend partial records. This method used as many weighted ratios as there were months in the monthly set to be used in making the estimate.

In general, the variances of the deviations were similar for these two methods, with errors of estimate of one method smaller for a given monthly set and the other more precise for another set. The single ratio was more precise for five of the eleven combinations in the 1-4-7-10 monthly set and for two of four combinations for each of the other two monthly sets. The single ratio method was more precise in all three sets when the full complement of months was included for that set; however, this superiority was small. Table 31 shows the errors of estimate of these methods to the standard deviation from regression with the single ratio being about two percent more efficient for the three sets of monthly data. Since 82.0 percent, 94.7 percent, and 93.3 percent of the animals are in milk ten, nine and eight months, respectively, logically more emphasis can be placed on the sets that include these months than for those sets containing lesser numbers of months.

TABLE 31

Relative Efficiency of Various Methods in Estimating Ten Month Production From Three Tri-Monthly Sets of Data

Method	Monthly sets		
	1(a)	2(b)	3(c)
Regression	100	100	100
Ratio	80.6 (d)	80.2	76.5
Inv. of Var.	77.3	78.2	74.0
DHIA	77.3	74.5	79.5

(a) set 1 is 1-4-7-10 months
(b) set 2 is 2-5-8 months
(c) set 3 is 3-6-9 months
(d) This is expressed as percent efficiency (E) compared to the regression estimate where E = 100 (residual variance of regression/residual variance of ratio).

Then the single ratio should be used in preference to the weighted ratios for extending the part records. However, to be able to project the ten-month total early in the laction is always important when either set in the preference is for the set that is easiest to obtain in a particular situation.

The equally weighted ratios provided the least precise estimates of any ratio method employed. There were only a few isolated instances where equally weighted ratios provided estimates as precise as either of the other two ratio schemes.

Regression

When multiple linear regression is used, the standard errors of estimates are lower than those of any other method. The precision of the estimates for any one monthly set is related to the amount of data available for that set, but some combinations involving the same number of variables are more precise than others in estimating the total production. The 1-4-10 monthly combination is the most precise of the three-variable sets in the 1-4-7-10 category. The 2-8 combination is most efficient of the two-variable sets from months 2-5-8 and the 3-9 set is the most precise of the three two-variable combinations in the 3-6-9 set.

It was postulated that regression factors and correlations derived from data where only records at least ten consecutive months in length were used would provide results inappropriate for the entire population. However, this was not the case when the regression factors derived from the ten-month data were applied to all cows in milk on test day. The residual variances were almost identical for both populations. This can be explained partially by the fact that 80 percent of those cows going dry prior to ten test dates went dry after the ninth test but prior to the tenth test date. Major discrepancies could occur only when the tenth test date alone was used to project a record since this was the month most cows were included in one group but not the other. But the lactation curve for the group going dry after the ninth test date was more like the curve for all cows in milk on last test date than for any of the other dry cow groups. For this reason, these records could not change the standard error of estimate very much. Differences of the other groups of dry cows could change the standard error, but their number was so few that it was not an effective force in changing the results.

The possibility of developing a separate set of regression factors to be used for extending the records of dry cows is not excluded. The standard error of estimate for dry cows only would certainly be quite high if one applied the factors developed for either the ten test day data or the group involving all cows. The use of separate regression factors is not very practical since just when a cow will go dry is not known at any given stage of the lactation. However to re-estimate the total lactation yield for dry cows after they go dry would be more accurate if any form of testing other than

monthly testing is being used.

Table 32 contains the efficiencies of the ten test date factors applied to all data. Relative efficiency is defined as the ratio of the error variance in the actual data to the error variance obtained when the 10 month data figures were applied to all data multiplied by 100.

Regression-Ratio Combinations

The ratio method is advantageous in that it does not include a herd effect. Regression factors can be developed which account for herd differences (VanVleck and Henderson, 1961d), but this method is rather cumbersome. The ratios provide a simpler relationship of the part to the whole, but these ratios when applied to all data lead to over and underestimates as noted by Harvey (1956). Thus, ratios are somewhat easier to use, butthe regression factors provide estimates closer to the actual production values. This led to the development of a set of regression factors that could be applied to the ratio estimates to provide an estimate of final production. This would allow ratios to be used for estimation in practical situations; but when more precise figures are required, the estimates could then be weighted and combined. These regression factors also give some insight to the precision of existing ratios. Table 33 shows the regression coefficients required to adjust the ratio estimate. For situations where the estimate by ratio was already highly correlated (.97) with the actual value, the addition of the regression factor did little to reduce the error of the estimate. Such was the case for the 1-4-7-10 estimate by ratio. The estimate by ratio was correlated with the actual value .98 and had a standard error of estimate of 16.7. When the additional step was added, the correlation remained the same and the standard error also remained nearly the same. The regression coefficient in this case was almost unity, .998. The facts differed when the estimate by ratio was from one month such as the first. Here the correlation between the estimate by ratio and actual value was .71 and the standard error of the estimate was 71.2; when the regression factor was used to extend this record, the correlation was still .71 but the standard error of estimate was reduced to 62.9. The regression coefficient used in this case was .65.

Distribution of Errors

All methods tend to overestimate the ten month production and the average estimate by regression method is nearer the actual average for the ratio combinations. The more information available the more precise is the estimate for both. On a 305 day basis the ratios overestimated 35 percent of the records by 610 lb of milk and underestimated 38 percent by the same amount. The proportionately weighted ratios estimated 71 percent of the observations within 619 lb of the actual production, and regression method estimated 73 percent of the observations within 543 lb of the actual production when the tri-monthly system of testing was employed, with

TABLE 32

Relative Efficiencies of Regression Factors Derived From Ten Month Data When Applied to all Cows in Milk on Test Data

Single Months	Standard Error of Applied Estimates (lb milk)	Standard Error of Actual Estimates (lb milk)	Relative Efficiency*
9	63.6	63.5	99.8
8	51.6	51.4	99.4
7	42.8	42.7	99.8
6	39.2	39.2	100.0
5	39.3	39.3	100.0
4	41.9	41.8	99.7
3	46.5	46.5	100.0
2	51.6	51.5	99.8
1	62.9	62.9	100.0
Sequential Months			
2-5-8	19.2	19.2	100
3-6-9	19.1	19.1	100

$$* \text{ Rel. Eff.} = \frac{\text{Resid. var. actual}}{\text{Resid. var. applied}} \times 100$$

TABLE 33

A Comparison of Efficiencies When Regression Factors Are Used to Extend Ratio Estimates to a Ten Month Basis

Monthly Set	Method	Regr. Coeff.[(a)]	R[(b)]	Std. Error[(c)]	% Eff.
1-4-7-10	Ratio (Cumulative)	.99	.99	16.7	100
1-4-7-10	Equally Weighted Ratios	.94	.95	24.7	100
1-4-7-10	Weighted Ratios (Proportionally)	.96	.97	17.1	100
2-5-8	Ratio	.95	.97	20.4	113
3-6-9	Ratio	.92	.97	20.0	119
1	Ratio	.65	.70	62.9	128
2	Ratio	.74	.81	51.5	122
3	Ratio	.77	.85	46.5	152
4	Ratio	.80	.88	42.0	123
5	Ratio	.80	.89	34.5	166
6	Ratio	.79	.89	39.2	133
7	Ratio	.73	.87	42.7	144
8	Ratio	.59	.80	51.5	185
9	Ratio	.42	.66	63.5	259
10	Ratio	.28	.53	71.5	340

$$\% \text{ Eff.} = \frac{\text{Resid. var. ratio est.}}{\text{Resid. var. ratio regr. Est.}} \times 100$$

a) Regr. Coeff. is the factor used to project the ratio estimate to a more accurate ten month basis.

b) R is the correlation coefficient between the new estimates and the actual values.

c) Std. Error is the residual variance of the new estimate.

a complete complement of months. This amounted to only 67 lb milk difference between the two extreme methods for about the same percent of the animals.

The ratios could be employed to estimate total production from part records with almost the same accuracy and be almost as precise as regression. All methods estimate more records within one and two standard deviations of the actual amounts than would be expected from normal distribution theory.

Application of "Ten Test Date" Regression Factors to All Data

The standard errors of estimate from regression factors where all cows are in milk at least ten consecutive test dates are smaller than corresponding values for the situation where all cows in milk on test data are included. While the former factors do not apply to the entire population quite as well as do those developed when all cows are included, they do provide more precise estimates than other methods. This would indicate that at the early stages of lactation, when which animals will go dry prior to 305 days or at what time are not yet known, the overall factors are best used to estimate production for ten months. However, at completion of the individual cow's lactation, an appropriate set of regression or ratio factors could be applied, the exact factors depending on the length of the record. The appropriateness of a method depends on the information desired. If an estimate is required early in the lactation, then the overall factorshould be used to provide the smallest number of errors.

If it is possible to wait until the record is complete, it may be more desirable to re-compute the ten month total from the appropriate set of regression factors. Since a large number of animals complete ten test dates, it would appear more feasible to begin by using the factors developed exclusively from those animals. Thus, 80 percent of the observations would already be predicted with the appropriate factors. The separate regression factors for various lengths of lactations have not been included in this study; however, ratios are available for this purpose (Aulerich, 1965).

APPLICATION OF RESULTS

Several methods of extending part records of various lengths have been presented. The regression factors are more precise in estimating the total lactation yield from part records on a cumulative, non-cumulative, or sequential basis than any of the ratio methods. The ratio methods, of which two were discussed in detail, are equally precise in most cases but both less precise than the regression factors. The more information available the more precise are the factors in estimating total yield from part records. The ratios, while not quite as precise as regression factors, are easier to compute and may be somewhat easier to use for practical situations than are the regression factors.

There are several purposes for using these factors in a production testing program. One is the application to a shorter interval testing program such as a bi-monthly or tri-monthly program. Another is to estimate production of animals at an early date in order to provide early information about sires for a sire proving program and particularly for a young sire program. The phenotypic correlations between parts and the whole indicate that partial records can be extended with considerable accuracy; however, Lamb (1962) reported that genetic progress per year will not be as rapid as if the complete records were used. This reduced accuracy may be offset if the generation interval is decreased sufficiently.

The cost of testing has been postulated to be a contributing factor to the relatively low percentage of herds enrolled in a production testing program. A less costly program, such as bi-monthly or tri-monthly testing could raise this percentage. To date, much information has been reported concerning the practical use of a bi-monthly program; however, most work has centered around the accuracy of the results only after the five tests have been reported. The use of either regression or ratio factors to extend a record to a ten-month basis each time the animal is tested may provide sufficient information to the participating herd owner to keep him interested in a testing program. Each time any additional data are available on the animal, a new, more accurate estimate can be computed. This not only will provide information which the herd owners can use each month in ranking their herds, but also these data will be available at all times for use in summarizing sires.

Most work available on bi-monthly tests assumes a full complement of monthly tests, but a substantial percentage of all cows terminate their records prior to having completed ten full months. For these situations and those where the animal is not tested for a particular month for some unknown reason, additional factors must be used. The results of this study provide regression factors for all combinations of data available from each of three monthly sets of a tri-monthly testing program. These factors and the resulting estimates must be included in the evaluation of a particular type of testing program if it is to be used in a practical manner.

Finally, ratios because of their simplicity may be more desirable for practical use. However, in order to provide a more precise estimate for sire summaries, to extend these estimates to a final "ratio-regression" estimate may be desirable.

SUMMARY

Data from 58,840 completed records of Michigan Holstein cows calving between January, 1959 and January, 1961 were used to evaluate several methods of estimating total lactation milk yield from various parts and to compare estimates obtained from several testing intervals with actual DHIA data.

Simple correlations were calculated between each of the various parts and between each part and total milk yield from only those records consisting of 10 consecutive test date data and from all records containing test date data through the last test date involved. Simple correlations were calculated between each of the various parts and between each part and total milk yield for each of 4 age-season groups.

Three ratio methods and linear regression were used to estimate total milk yield from the different parts. The non-cumulative ratio was used to estimate total yield from each part, and these estimates were weighted equally or weighted according to the variance of each individual estimate to form pooled estimates. The third ratio method involved the formation of a single cumulative ratio from data in the non-cumulative months. Total DHIA yield was obtained after summing data for each cow over all test dates.

Total milk yield was estimated from each set of data obtained by two different bi-monthly testing plans and three different tri-monthly testing plans, as well as from all possible partial sets of data for each plan. Total milk yield was also calculated by the linear regression of sequential, cumulative, and non-cumulative yield on time.

Regression factors for each case were obtained for each of 4 age-season groups.

Regression factors to estimate total milk yield were computed from only those records consisting of 10 consecutive test date data and from all records containing test date data through the last concerned test date.

Simple linear regression factors were obtained to estimate total milk yield from the ratio estimates for each set of bi-monthly and tri-monthly data.

Correlations between single monthly parts and total milk yield were highest for the middle months (5 and 6) and lower for either extreme from these months, with the tenth month being the lowest correlated single month with total yield (0.53). Correlations were higher between each single monthly part and total milk yield for data including only ten consecutive test dates than for data including all cows in milk on test date. The greatest differences occurred at the early months of lactation and converged as the monthly number increased. Simple correlations between various parts were highest for adjacent months, again, the data restricted to ten consecutive test dates were higher correlated than the unrestricted data.

Estimates obtained by the equally weighted ratios were less precise than estimates from either of the other ratio methods. The latter estimates were equally precise, but were less precise than estimates obtained by linear regression. Estimates obtained from the more precise ratio method were within 509,653, and 655 lb milk of actual values for tri-monthly sets consisting of months 1-4-7-10, 2-5-8, and 3-6-9, respectively. Bi-monthly estimates from the same ratio methods were within 144 and 135 lb. milk of actual values for bi-monthly sets consisting of months 1-3-5-7-9 and 2-4-6-8-10, respectively.

Means of total production were highest for animals initiating lactations at 36 or more months of age and from August-March and were lowest for animals calving during April-July at ages less than 36 months. The younger cows reached peak production at lower levels and did not diminish production as rapidly as the older animals. Larger regression factors were required to estimate total milk yield for younger cows from cumulative production in the early months of lactation and become about equal during the middle and latter months of lactation.

Regression factors to estimate total milk yield from various parts are more precise when obtained from data including only those animals in milk ten consecutive test dates.

However, all cows will not be in milk the full ten months of lactation thus, it may be desirable to recompute the total milk yield using the appropriate regression factors after the cow has gone dry.

Application of simple linear regression to the ratio estimates provides a more precise estimate of total milk yield than does the ratio estimate alone. The efficiency of doing this in terms of the residual variance decreases as additional monthly data become available. Efficiency values range from 340% when month one alone is used to predict total milk yield to 100% for tri-monthly sets of data consisting of months 1-4-7-10. Comparable results for monthly sets 2-5-8 and 3-6-9 are 113% and 119%, respectively.

LITERATURE CITED

Alexander, M.H. and Yapp, W.W. (1949). Comparison of Methods of Estimating Milk and Fat Production in Dairy Cows. J. Dairy Sci., 32:621.

Aulerich, C. (1965). Differences in Extending Terminal and Non-Terminal Incomplete Lactations to 305 Days. Unpublished M.S. Thesis, Michigan State University Library, East Lansing.

Bayley, N.D., Liss, R.M. and Stallard, J.E. (1952) A Comparison of Bi-Monthly and Quarterly Testing with Monthly Testing for Estimating Dairy Cattle Production J. Dairy Sci., 35:350.

Campbell, I.L. (1946 and 1948). The Accuracy of Various Methods of Estimating a Dairy Cow's Production of Milk and Butterfat. Proc. Ann. Conf. New Zealand, Soc. Anim. Prod. 99-104. Dairy Sci., Abs., 9:258.

Cannon, C.Y., Frye, J.B., Jr. and Sims, J.A. (1924). Predicting 305-Day Yields from Short-Time Records. J. Dairy Sci., 25:991.

Copeland, L. (1928). Monthly and Bi-Monthly Tests. Jersey Bul. 47:731.

Cunningham, E.D. (1965). The Relative Accuracy of Different Sampling Intervals and Methods of Prediction in Estimating Milk Yield and Composition. The Agric. Institute - Dublin (Abs.)

DHIA Newsletter. DHIA Factors for Projecting Incomplete Records to 305 Days. Aug., 1965.

Dick, I.D. (1950). Some Results on the Accuracy of Milk Sampling Over a Complete Lactation and the Daily Variability in Milk Yields. New Zealand Sci. Technology; 32(A):25-29.

Erb, R.E., Goodwin, Mary M., McCaw, W.N., Morrison, R.A. and Shaw, A.C. (1953). Lactation Studies VI. Improving the Accuracy of Larger Testing Interval and the Accuracy of Current Methods. Wash. Agr. Expt. Sta. Circ. 230.

Erb, R.E., Goodwin, Mary M., Morrison, R.A. and Shaw, A.O. (1952). Lactation Studies IV. Accuracy of Different Methods of Estimating Lactation Yields. J. Dairy Sci. 35:977.

Flanagan, M.J. (1955). The Accuracy of Various Methods of Recording in Estimating Actual 300 Day Milk Yield in Dairy Cows. Unpublished M.S. Thesis. National University of Ireland.

Fritz, G.R., McGilliard, L.D., Madden, D.E. (1960). Environmental Influences on Regression Factors for Estimating 305 Day Production from Part Lactation. J. Dairy Sci., 43:1108.

Gaines, W.L. (1927). The Deferred Short-Time Test as a Measure of the Performance of Dairy Cows. J. Agri. Red., 35:237.

Gifford, W. (1930). The Reliability of Bi-Monthly Tests. J. Dairy Sci., 13:81.

________. (1943). The Value of Partial and Complete Lactation Records for Evaluating Dairy Cows. J. Dairy Sci., 26:724 (Abs.).

Gowen, M.W. and Gowen, J.W. (1922). Studies in Milk Secretion XVII. Relation Between Milk Yields and Butterfat Percentages of the 7-Day and 365 Day Tests of Holstein-Friesian Advanced Registry Cattle. Maine Agri. Sta. Bull. 306.

Harvey, W.R. (1956). Extension of Incomplete Records to a 10-Month Basis. Unpublished Data from Ayrshire Herd Test Records.

Johansson, E. and Hansson, A. (1940). Causes of Variation in Milk and Butterfat Yield of Dairy Cows. Kungl. Lantbr. Tidskr. 79(6½):1.

Kennedy, C.M. and Seath, D.M. (1942). The Value of Short-Time Records for Culling and for Progeny Testing of Dairy Cattle. J. Animal Sci., 1:348.

Lamb, R.C. (1959). Variables Affecting Ratio Factors for Estimating 305 Day Production from Part Lactation. MS. Thesis. Michigan State University Library.

________. (1962). Usefulness of Part Records in Estimating the Breeding Value of Dairy Cows. Ph.D. Thesis. Michigan State University Library.

Madden, D.E., Lush, J.S., McGilliard, L.D. (1955). Relations Between Parts of Lactation and Producing Ability of Holstein Cows. J. Dairy Sci., 38:1264.

_________, McGilliard, L.D. and Ralston, N.P. (1956). Relations Between Monthly Test Day Milk Production and Holstein Cows. J. Dairy Sci., 39:932 (Abs.)

_________, McGilliard, L.D. and Ralston, N.P. (1959). Relations Between Test Day Milk Production of Holstein Dairy Cows. J. Dairy Sci., 42:319.

McCandlish, A.C. and M'Vicar, A. (1925). Are Milk Record Association Results Accurate? Scot. J. Agri., 8: 201-206.

McDowell, J.C. (1927). Testing Cows for Production Every Other Month. Holstein-Friesian World, 24:1970.

McKellip, I., and Seath, D.M. (1941). A Comparison of Different Methods of Calculating Yearly Milk and Butterfat Records. J. Dairy Sci., 24:181.

Miller, C.C. (1963). Unpublished Data. Dairy Dept., Michigan State University, East Lansing.

O'Connor, L.K. and Lipton, S. (1960). The Effect of Various Sampling Intervals on the Estimation of Lactation Milk Yield and Composition. J. Dairy Res. 27:389.

Rabild, H. (1909). Cow Testing Associations. Ann. Rpt. Bur. Animal Ind., U.S.D.A. 99-118.

Reece, R.P. (1942). Can One Predict the Average Fat Test for a Lactation on the Basis of an Incomplete Record. J. Animal Sci. 1:349.

Rendel, J.M., Robertson, A., Asker, A.A., Khishin, S.S. and Regab, M.T. (1957). The Inheritance of Milk Production Characteristics. J. Agri. Sci., 48:426.

Turner, C.W. (1926). A Quantitative Form of Expressing Persistency of Milk or Fat Secretion. J. Dairy Sci., 9:203.

_________, and Ragsdale, A.C. (1924). A Comparison of Holstein-Friesian Sires. Mo. Agri. Expt. Sta. Bull. 217.

VanVleck, L.D. and Henderson, C.R. (1961a). Regression Factors for Estimating Part Lactation Milk Records. J. Dairy Sci., 44:1085.

________, and ________. (1961b). Ratio Factors for Adjusting Monthly Test-Day Data for Age and Season of Calving and Ratio Factors for Extending Part Lactation Records. J. Dairy Sci., 44:1093.

________, and ________. (1961c). Regression Factors for Predicting a Succeeding Complete Lactation Milk Record from Part Records. J. Dairy Sci., 44:1322.

________, and ________. (1961d). Extending Part Lactation Milk Records by Regression Ignoring Herd Effects. J. Dairy Sci., 44:1519.

Voelker, H.H. (1957). Use of Extended Incomplete Lactation Records, J. Dairy Sci., 40:651 (Abs.).

Yapp, W.W. (1915). Relative Reliability of Official Tests for Dairy Cows. Ill. Agri. Expt. Sta. Bull. 215.