ENVIRONMENTAL INFLUENCES ON REGRESSION FACTORS FOR ESTIMATING 305-DAY PRODUCTION FROM PART LACTATIONS

Thesis for the Degree of M. S.

MICHIGAN STATE UNIVERSITY

George R. Fritz, Jr.

1958

ENVIRONMENTAL INFLUENCES ON REGRESSION FACTORS FOR ESTIMATING 305-DAY PRODUCTION FROM PART LACTATIONS

By

George R. Fritz, Jr.

AN ABSTRACT

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

MASTER OF SCIENCE

Department of Dairy

Year 1958

Approved Lon D. McGilliard

Cumulative milk and fat production on test day from 11,420 Michigan D.H.I.A. - IBM records for the period 1953 to July 1957 were analysed, and used to derive regression factors for estimating 305-day production from part lactations. The data included 8,993 Holstein, 1,457 Guernsey, 651 Jersey, and 319 Brown Swiss records.

Linear regression coefficients were used as the variable measuring the relationship between cumulative production on test day and the sum of the first ten test days. The regression coefficients were calculated within each herd, season, lactation, and age group. Analyses of variance and visual inspection of the milk and fat regression coefficients for the first, third, and seventh cumulative months of the Holstein data were used to determine the importance of season, lactation, and age on the part to whole relationship. From these techniques, season was not considered an important influence for the present study. Visual inspection of the mean age coefficients within lactation numbers failed to reveal where the groups should be separated or combined; consequently, the data were grouped as a whole within each breed.

Intra-herd extension factors were computed as the weighted average regression coefficient within herds, seasons, lactations, and ages for each of the nine cumulative test days. Each regression coefficient included in the average was weighted by the number of records determining its value. For comparison, inter-herd extension factors were also derived disregarding herd, season, lactation, and age effects. Although the inter and intra-herd factors were not tested for significance, the marked similarity between the

ABSTRACT GEORGE R. FRITZ, JR.

Holstein factors suggests that herd has little or no influence on the part to whole relationship.

Correlations between cumulative part and 305-day production ignoring herd, season, lactation, and age differences were not less than .7 for the first month, increased steadily as the lactation progressed; and were .9 by the fifth test day for all breeds. These findings support data in the literature that suggests production records of only one or two months are valuable guides to what a cow will produce in that lactation. Furthermore, estimating 305-day production from such records would be useful tools in early culling and progeny testing.

ENVIRONMENTAL INFLUENCES ON REGRESSION FACTORS FOR ESTIMATING 305-DAY PRODUCTION FROM PART LACTATIONS

By

George R. Fritz, Jr.

A THESIS

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

MASTER OF SCIENCE

Department of Dairy

Year 1958

1-16-59
[illegible]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The author wishes to thank Dr. D. E. Madden for suggesting the research topic and his patient guidance throughout the investigation.

Appreciation is also extended to Dr. L. D. McGilliard for his friendly council during the investigation and helpful suggestions in preparation of the thesis.

Thanks are expressed to Dr. N. P. Ralston, Head of the Dairy Department, for teaching and research assistantships affording the opportunity for graduate study and valuable teaching experience.

To Mr. Robert P. Witte for his invaluable assistance in formulating the machine procedures, and to Mr. Robert C. Lamb for his aid in processing portions of the data go sincere thanks.

Most of all, the author is ever indebted to his family and Miss Margaret Rood for their continued and buoyant encouragement.

INTRODUCTION

Approximately half of the cows leaving the milking herd before completing their 305-day records are removed because of low production. For the three year period 1955-57, the number of Michigan Dairy Herd Improvement Association (D.H.I.A.) - IBM records not completed for this reason was 5,899 (11, 12, 13). This constituted a loss of production records equal to eleven per cent of the 53,286 completed records for the same period.

Excluding such a large proportion of records which are incomplete primarily due to low production distorts tremendously the results of sire proofs. Eliminating low producing daughters before they can complete their first production record is a serious type of record selection which introduces a severe bias in sire proofs. Sires appear better than they would if all records were included. If it were possible to estimate with reasonable accuracy 305-day production from what partial production information is available, incomplete lactation records could be included in sire proofs. This would eliminate selective progeny testing resulting from the removal of low producing cows before they can complete their first record of production.

The large correlations found between the first and subsequent lactation yields indicate that the first lactation is a good measure of what a cow will produce in future lactations. Extending incomplete records could therefore provide an early evaluation of bulls and their daughters before the daughters have completed their first records of production. Culling decisions on cows early in their first or second lactation would also be possible by extending incomplete records.

Several workers have reported the ability to estimate the complete lactation yield by applying multiplicative and regression factors to an incomplete record or to production tests made during certain portions of the lactation. These techniques assumed that similar relationships between production during different portions of the lactation and the total yield exist for all cows.

Previous work has indicated that (1) breed, (2) age and (3) season of freshening, (4) milking frequency, and (5) stage of gestation influence the quality and/or quantity of milk and fat produced during all or certain portions of the lactation. Whether these factors exert an influence on the relationship between an incomplete record and its 305-day equivalent has not been determined. Recently, several workers have delved into this problem with limited success by studying the differences between extension factors computed for different seasons, ages of freshening, and milking frequencies. To assure maximum accuracy of the estimated 305-day production, all environmental agents influencing the part to whole relationship must be compensated for in the method used in deriving the predicting factors.

The present study was initiated to (1) examine the importance of breed, herd, lactation number, season, and age of freshening on the relationship between milk and fat produced on test day and production for the complete lactation, and (2) to compute factors for extending incomplete records to a 305-day basis.

REVIEW OF LITERATURE

Voelker (1957) found that sixteen per cent of all production records available (1,636) in the South Dakota College herd since 1897 were incomplete due primarily to low production. The per cent of records terminating in the first to the ninth month, respectively, were 5, 11, 9, 11, 13, 15, 18, 12, and 6.

Two methods of deriving factors to estimate 305-day production from yields of less than 305-days have been used in the past. The simplest procedure is the ratio of the complete production to partial production. It has a general predicting equation of the following form: $\hat{Y} = cX$, where $\hat{Y}$ is estimated production for 305-ddys, c the ratio of complete production to partial production, and X the actual production for less than 305-days.

The second technique is to fit by "least squares" a regression equation of total production on partial production. The regression procedure has a predicting equation of at least two parameters and is somewhat more complex in application. In its simplest form (linear regression), the general equation has the structure: $\hat{Y} = a + bX$, where $\hat{Y}$ is estimated production for 305-days, a the point where the regression line intercepts the y axis, b the regression coefficient measuring the average change in Y with each unit change in X, and X the actual production for less than 305-days.

Unless a regression of total production on month of production utilizing all nine months is used as a basis of predicting total production, it is necessary to have different equations for each month regardless of the method (ratio or regression) since cows do not produce the same quantity of milk and fat throughout the lactation. Consequently, a, b, and c in the two general equations will take on different values depending on the month or length of lactation from which the total yield is being estimated.

Both methods will extend an incomplete record, but the regression equation adjusts for the incomplete repeatability of different portions of the lactation at the same time. Unless some adjustment is made after the incomplete record has been extended by the ratio equation, estimates of total production by the two methods would vary for the same lactation. The difference between the estimates would be the amount that must be added to cX to equal $a + bX$. This difference is $(b - c)(X - \bar{X})$ where b represents the linear regression of total on part, c the ratio of total to part, and X and $\bar{X}$ represent part and average part production. The difference $(b - c)$ is negative and largest $(c > b)$ during the early (1-3) cumulative months, least $(c \sim b)$ during the center (4-6), and positive and small during the longer (7-9) cumulative parts (Madden et al. 1959). Total production estimated by the ratio method would have a tendency to overestimate for the high producers $(X > \bar{X})$ and underestimate for the low producers $(X < \bar{X})$ during the first 3-4 months of the lactation. The increased variability of estimated 10-month production from records of one, two, and three months duration was shown to be 12-47 per cent by the ratio method as compared to the linear regression equation (Harvey 1956).

Gaines (1927) examined milk and fat percentage data from the Guernsey and Holstein Advanced Registries. He found that a one or two day test conducted between the fourth and seventh month of the lactation best represented a cow's average total production for twelve months. This conclusion was based on a study of the correlations between the monthly and yearly fat percentages which were the largest for the four middle months, ranging from .76 - .80 and .82 - .83 for the one and two day tests, respectively.

A similar conclusion was drawn by Cannon et al. (1942) while working with 400 Holstein records from the Iowa State College herd and 1,289 D.H.I.A. lactation reports. Predicting factors (regression) for each month of a 10-month period were calculated from butterfat yield for a single day for the two groups of data separately. Production predicted for ten months from the fifth month was considered the most accurate since it had the lowest standard error of estimate, 43.8 and 45.6 for the college and D.H.I.A. factors, respectively. Tests during the sixth month were the next most accurate followed closely by the seventh and fourth. These findings of Gaines and Cannon indicated that there is less accuracy in predicting a cow's lactation yield from a single test in the initial and final months than there is from a test in the middle portion of the lactation.

In view of Gaines' and Cannon's results, it seems logical that the first 4-5 consecutive months production couldn't be less accurate than a test of production for a single day as a means of predicting total production. Kennedy and Seath's investigation (1942) of the value of incomplete records as a basis for culling and progeny testing supports this proposition. They found correlations between short-time and complete records of 305-days duration increased from an average of .62 for the first to .78 by the end of the first four months of the first lactation. Records of 80 Holsteins and 80 Jerseys having complete records for their first two lactations in the Louisiana State University herd were used for their analysis. Rendel _et al_. (1957) discovered that milk yields as short as the first 70 days of the lactation were excellent guides to predicting total production. High correlations (.8) between records of that length and total production were found in 3,109 production records of six main dairy breeds in England and Wales. When Voelker (1957) extended current two-year-old records in the South Dakota College herd, the correlations between the actual 305-day and estimated yield from 1-9 months production were .68, .74, .76, .85, .89, .92, .94, .95, and .99, respectively.

Cannon _et al_. (1942) also computed ratio factors to estimate what a cow would produce when only a portion of the lactation record was known. The predicting factors ranged in magnitude from 7.47 at the first month to almost unity at the ninth. Differences were noted between the college and D.H.I.A. values, especially during the first four months. This was probably due to differences in management and not separating the D.H.I.A. data by breeds.

Gifford (1939) made an intensive study of correlations between individual months of consecutive lactations with 2,600 pairs of records from 99 herds in Iowa Cow Testing Associations. He found that each month was more closely correlated with the production of the other months in the same lactation than with the production for the same month in a subsequent lactation. This suggests that there are certain environmental factors influencing production during all or nearly all portions of a specific lactation which are not constant from one lactation to the next. The study also indicated that short-time records of only a few months had a high value as indicators of a cow's real producing ability. Real producing ability is a cow's capacity to produce under normal conditions.

The large genetic correlations (.9 or larger) between cumulative part and complete milk and fat production found by Madden et al. (1955) demonstrated that genetic factors influencing one portion of the lactation also influenced all others as well. These results were derived from 599 records completed by 253 Holsteins in the Iowa State College herd from 1940-52. Madden also studied the influence of age at freshening on the part to whole relationship by comparing ratio factors computed for different ages from actual cumulative production. Two age groups were used, records initiated before three years of age and those started after three years. This grouping procedure was determined from plotting average monthly production for the various ages of freshening and noting a distinct difference between the lactation curves of the cows under three years as compared to those of cows over three years of age. Because the extension factors were so dissimilar in the initial months, it was concluded that separate factors were necessary for the two age groups at least for the first 150 days. The extrapolation factors for milk for the initial 30 days were 9.36 and 7.81 for the ages under three and over three years, respectively. The differences between the two sets of factors diminished as the lactation advanced. These findings demonstrated that young cows start producing at a lower level than the older cows but do not decrease in production as fast in the latter half of the lactation. Similar results regarding young cows were noted in a later paper by Madden et al. (1956) when ratio factors were obtained from 6,495 Holstein H.I.R. records.

Kendrick (1955) investigated the influence of season and level of production in addition to age at freshening by compiling ratio factors from 736 Ayrshire Herd Test records. Factors were compared for three age groups (under 30 months, 31-44 months, and 45 months and older), two seasons (April-September, and October-March) and two levels of production. The production levels were determined by the first 120 days production of each lactation record. Below average production was designated "low" and above average "high." The data supported findings of Madden et al. (1955) that separate factors would be necessary for different ages at freshening because of the larger factors found for the younger cows. After the sixth month, the factors developed apparently were not affected by season of freshening or level of production.

As a supplement to Kendrick's sample of Ayrshire data, Harvey (1956) obtained an additional 2,131 Herd Test records for statistical analysis. The combined 2,867 Ayrshire records were separately analysed for three age groups, cows calving at 32 months or less, between 33-46 months, and 47 months of age or older. Linear and quadratic equations were fitted to the accumulated milk and fat production data for the nine consecutive test days. From the small differences between the multiple and simple correlations, Harvey concluded that the curvilinearity of the relationship between part to whole accumulated production was small when averaged over all herds. It is possible that herd differences have masked the true picture.

In a more recent paper, Madden et al. (1959) studied further the relationships between production for a single or cumulative test day and

305-day production. Age at calving, milking frequency and level of production were considered as having possible influence on how extension factors should be developed. The 6,715 Holstein H.I.R. records studied were grouped according to milking frequencies (2x and 3x), and age groups previously used in the study at Iowa (Madden et al. 1955). Linear and quadratic regression equations were fitted to the data to determine the influence of level of production on the part to whole relationship. The small differences found between the simple and multiple correlations indicated that linear regression adequately described the relationship of part to whole cumulative production for high as well as low producing cows. No significant differences were found between extension factors for 2x and 3x frequencies. This suggests that a similar relationship is prevalent between the incomplete and complete 305-day production regardless of the number of times the cow is milked per day.

MATERIALS AND PROCEDURE

SOURCE OF DATA

An estimated six or seven hundred thousand Michigan D.H.I.A. - IBM monthly test day production cards for the period 1953 to July 1957 were available for use in this study. The IBM form of the D.H.I.A. testing program processes by machine the information gathered by the supervisor.

Each record identified the (1) cow, (2) herd, (3) date of calving, (4) age of calving, (5) lactation number, and (6) breed. Production on test day was determined by the D.H.I.A. supervisor weighing and sampling the total milk produced for a single day once each month and testing it for per cent of fat. Total milk production on test day for each cow was recorded to the nearest one tenth of a pound. The amount of fat on test day to the nearest one hundredth of a pound was obtained by multiplying the pounds of milk times the per cent of fat.

The first ten consecutive monthly test days of all lactations having ten or more test days were gathered from the data. Of these records, lactations having more than 49 days elapsing between the date of freshening and the first test were excluded from the study along with lactations having less than ten consecutive monthly test days. This removed possibilities of having ten consecutive monthly test days of a lactation but the cow not being tested the first month after calving. Only four breeds had sufficient numbers of records for analysis. The number of records for the breeds were: 12,562 Holstein, 2,262 Guernsey, 990 Jersey, and 459 Brown Swiss.

Since large correlations have been found between an incomplete record of only a few months and the total yield (Rendel *et al.* 1957) and "total-to-date" is reported in most testing programs, production accumulated for successive test days instead of production for individual test day was used to predict 305-day production. By using accumulated production, the effect of any environmental factor influencing the yield of any one test day would be diluted and have an opportunity to be cancelled if the influence were random from month to month. Production on test day is converted to cumulative-to-date production by multiplying by 30.5, the average number of days in a month.

Lactation number as well as herd, season, and age differences were considered as possible important influences on how extension factors should be computed. From previous studies with Ayrshire Herd Test records (Kendrick 1955; Harvey 1956) and Holstein Herd Improvement Records (Madden *et al.* 1959), it was suggested that separate extension factors should be developed for different seasons and ages of freshening. This was considered necessary because of climatic changes and production characteristics of the younger cows reflected a different part to whole relationship when factors were developed and compared for different seasons and ages.

The question, can partial records of cows freshening for the second time be accurately extended if treated in the same manner as those freshening for the first time at the same age, has not been examined previously. The first, second, and third lactations of the Guernsey and Holstein data were plotted separately with age of calving along the abscissa and number of records along the ordinate in Figures 1 and 2.

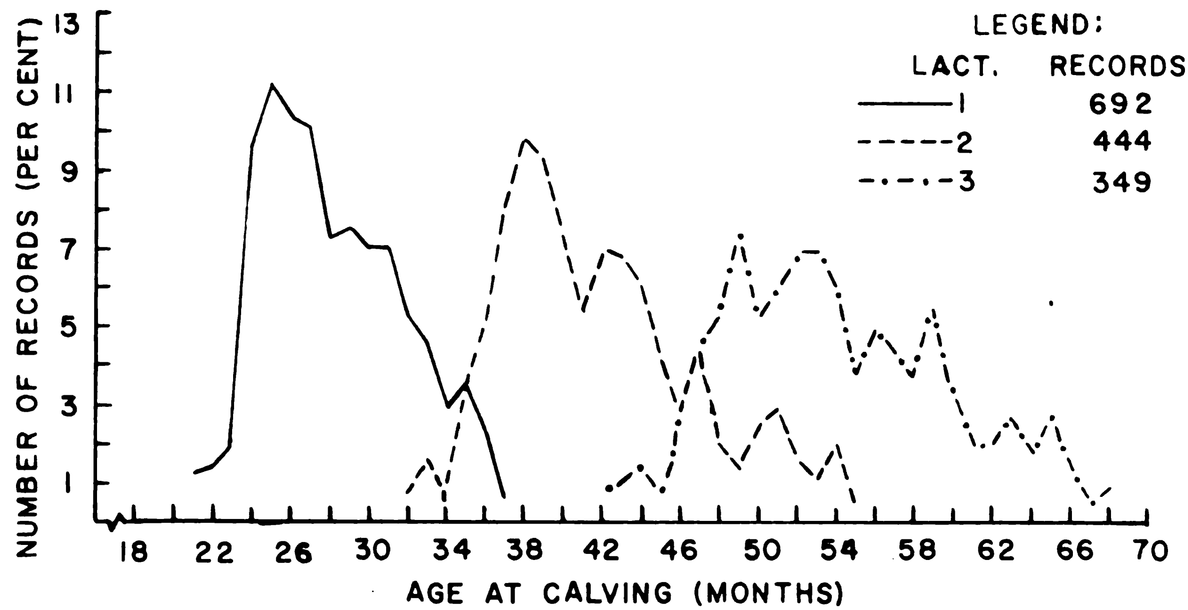

FIG. 1 FREQUENCY OF GUERNSEY RECORDS BY AGE AND LACTATION NUMBER

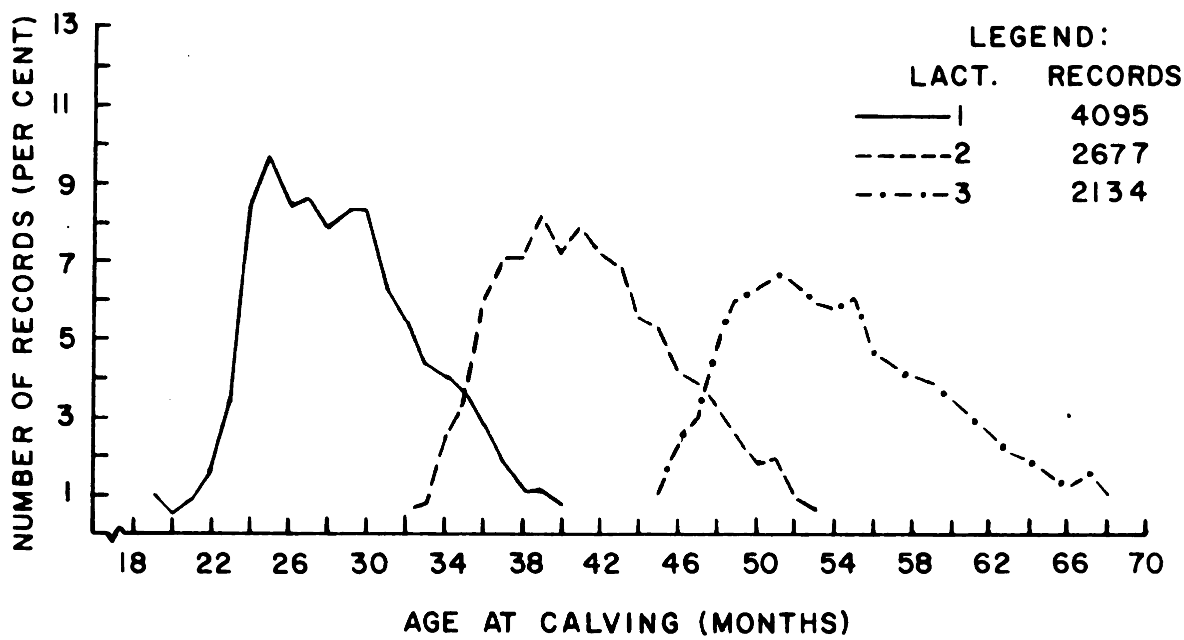

FIG. 2 FREQUENCY OF HOLSTEIN RECORDS BY AGE AND LACTATION NUMBER

The first lactation curve intercepts the second lactation curve at 35 months of age. The second lactation curve intercepts the third at 47-48 months. Similar characteristics were observed by Madden et al. (1959) with Holstein data. They assumed since about proportional percentages of first lactations were initiated after 36 months (10%) as second lactations prior to 36 months (12%) and the most important production information was lost in first lactations, a division of ages into under-three-years and over-three-years was sufficient. To answer the question of the effect of early and late first calving on the part to whole relationship, it was necessary to make further refinements in grouping the data by ages.

In order to consider the postulated environmental influences, the data were grouped by two 6-month seasons of freshening, three lactation numbers, and eight ages. The two seasons were used to measure a management system of dry roughage feeding (October-March) and pasturing (April-September) as well as climatic conditions in Michigan. A similar grouping was used by Kendrick (1955).

Since the need for extending first lactations is the most important, it would be logical to group them separately, but there was no assurance that all subsequent lactations could be treated as a single group. Consequently, the first two lactations were individually separated from all others. This grouping also facilitated determining if possible the significance of early and late first calving on the part to whole relationship. The majority of records falling in the third lactation group would likely be from mature cows if the records were plotted by age.

The age groups were divided by 6-month intervals, but all records initiated at less than 24 months were in a single group and all 60 months or over in another.

<u>MEASURE OF RELATIONSHIP</u>

Linear regression coefficients within herds, seasons, lactations, and ages were used as the variable measuring the relationship between the production for the cumulative part and the complete ten test days. All records in the same herd having the same lactation number, season and age of freshening would be exposed to similar environmental conditions. Because the hereditary relationship between the part and whole production has been found to be close to 1 (Madden *et al.* 1955), the hereditary differences in production should not change the part to whole relationship. Any variation between the intra-herd coefficients for the various season, lactation number, and age groups would reflect the significance of the environmental agents under consideration. Although it was a more cumbersome statistic with which to work, the regression coefficients were used instead of ratios because it was decided to develop regression factors after determining which environmental agents were important. Regression factors will adjust for incomplete repeatability of different portions of the lactation at the same time the incomplete record is extended.

At least two observations were necessary to fit a regression line by "least squares" to each intra-herd, season, lactation, age group; therefore, all records not meeting the requirement were eliminated from the study. The number of usable records remaining within the four breeds were 8,993 Holstein, 1,457 Guernsey, 651 Jersey, and 319 Brown Swiss. The linear regression coefficients were then calculated to the nearest one hundredth of a pound for each intra-herd, -season, -lactation, -age group separately for the nine cumulative test days.

DETERMINING INFLUENCE OF SEASON

To acertain if season of freshening should be considered in computing predicting factors, analyses of variance were performed on the first, third, and seventh cumulative test days of the Holstein data. The regression coefficient within each herd, season, lactation, and age group was used as the observed variable. Only three test days were analysed because previous investigations (Harvey 1956; Madden *et al.* 1959) had shown the correlations between cumulative part and the complete lactation were least for the first month and increased progressively as the lactation continued. These correlations were all .9 by the fourth month. The three test days analysed were the most convenient to process and yet marked the low, middle, and high correlations previously found.

DETERMINING INFLUENCE OF AGE AND LACTATION

To determine how ages and lactation numbers should be grouped or separated, unweighted means of the intra-herd regression coefficients for ages within the three lactation groups were scrutinized for the same three test days utilized in the analysis of season.

COMPUTING EXTENSION FACTORS

Extension factors were computed as the average regression coefficient within herds, seasons, lactations, and ages for each cumulative test day. The regression coefficients derived fluctuated quite radically from one month to the next in all breeds except Holsteins. Weighting each coefficient in the average by the number of records included reduced the variation from month to month, and the weighted average coefficient was taken as the extension factor.

RESULTS AND DISCUSSION

MEASURE OF RELATIONSHIP

Linear regression coefficients within herds, seasons, lactation numbers, and ages were utilized to measure the relationship between production during parts of the lactation and complete production for ten test days. Young cows begin their lactation at a relatively low level of production and decline gradually after reaching their peak production. In contrast, older cows have a higher initial and peak production and decline quite rapidly as the lactation advances. The question naturally arises is the part to whole relationship on a cumulative production basis more accurately described by curvilinear regression rather than linear regression. Both Harvey (1956) and Madden et al. (1959) have demonstrated that the curvilinearity of cumulative part to whole production was slight when averaged over all herds. This was done by fitting linear and quadratic equations to the data separately and comparing the simple and multiple correlations. The differences were found to be small for both Ayrshire and Holstein records. The curvilinearity of cumulative part to whole production on an intra-herd basis of the present data was not determined. Linearity was assumed for the present investigation.

The regression coefficients within herd, season, lactation, and age for total production on the amount for the first test day assumed a wide range of positive and negative quantities. But, as additional production information was added in succeeding test periods, the variation in the regression coefficients for subsequent test days decreased and the frequency of negative values was reduced.

INFLUENCE OF SEASON

The structure of the analysis of variance prepared to determine the importance of season on the relationship between the part and complete lactation is presented in Table 1. The significance of the analysis is weakened by all herds not having proportional numbers of cows freshening in both seasons. Consequently, if the Mean Square for interaction (Herd X Season) is used to test the significance of the variation attributed to herds and seasons, the probabilities will be inexact and indicate trends rather than precise differences. An analysis of a proportional group of observations would be more conclusive in answering the question of seasonal influence on the part to whole relationship, but these do not occur.

Variation due to season was significant for the first test day of fat production exclusively. Highly significant herd differences were noted for the first test day of milk production. Herd, season interactions were highly significant for the first test day of fat, and for the seven cumulated test days of milk and fat. Significant differences in fat production due to season; especially for the first test day, would be expected since a direct relationship of season of the year and per cent of butterfat has been found regardless of the stage of lactation (Becker and Arnold 1935; Wylie 1923). The lowest fat test usually occurs during June-July and the peak test in December-January. Because the smallest percentage of incomplete records are terminated in the first month of the lactation (Voelker 1957) and the seasonal effect for the first month is expected, season was not considered an important environmental factor

Table 1. ANALYSIS OF VARIATION IN REGRESSION COEFFICIENTS FOR HERDS AND SEASONS - Holstein

Test Day	Source	DF	Milk Mean Square	Fat Mean Square
1	Total	2698		
	Between Seasons	1	476	986*
	Between Herds	634	1417**	255
	Herd X Season [a]	421	1005	253**
	Residual	1642	1208	215
3	Total	2698		
	Between Seasons	1	00	21
	Between Herds	634	98	60
	Herd X Season [a]	421	77	59
	Residual	1642	114	80
7	Total	2698		
	Between Seasons	1	30	49
	Between Herds	634	25	11
	Herd X Season [a]	421	39**	16**
	Residual	1642	13	12

* P < .05   **P < .01 due to disproportionality these are approximations

[a] 94 herds had records in the Oct.-Mar. season exclusively; 112 herds had records in the Apr.-Sept. season exclusively.

influencing the relationship between the part and whole lactation.

INFLUENCE OF AGE AND LACTATION

The unweighted average regression coefficients of total on part lactation for cows of the same age in the same lactation for the Holstein data are presented in Table 2. No distinct pattern was observed where ages differed within or between lactation groups. The number of coefficients determining averages for the 30-35 and 36-41 age groups of the first and second lactations were the most nearly proportional. Yet the relationship between the age coefficients for a given month was not consistent for the three days analysed.

A comparison of records initiated before and after 36 months of age for the first and second lactations showed less overlapping of the two lactations than reported by Madden et al. (1959). Only 7.0 per cent of the first lactation records were initiated after 36 months and 5.2 per cent prior to 36 months for the second lactations. Upon these findings, age and lactation number were ignored and regressions were averaged over all ages and lactations within each breed.

Probably a more adequate evaluation of the age, lactation influences could have been obtained with an analysis of variance by testing the differences between lactation numbers having records in the same age groups. Following this the data could be regrouped accordingly; i.e. if there was no difference between lactation numbers for the same ages, the lactations could be combined. Then the ages could be tested with Duncan's Multiple Range Test to determine where they should be separated.

UNWEIGHTED EXTENSION FACTORS

Extension factors based on cumulative production would be expected

TABLE 2. UNWEIGHTED AVERAGE REGRESSION COEFFICIENTS FOR AGES AND LACTATION GROUPS - Holstein

Test Day	Lact. Group	Age Group -24	24-29	30-35	36-41	42-47	48-53	54-59	60/
		(No. of Coefficients)							
	1	53	493	318	67	12	--	--	--
	2	--	1	33	284	206	48	9	4
	3	--	--	1	5	27	193	176	769
		(Lbs. of Milk)							
	1	5.55	6.63	6.33	4.05	6.89	--	--	--
1	2	---	-.61	5.35	4.90	4.17	6.46	14.62	-79.44
	3	---	---	27.54	20.98	4.20	6.27	4.82	4.75
	1	-.49	1.89	2.93	2.56	2.60	--	--	--
3	2	---	-.56	3.47	2.34	2.27	1.51	2.04	4.09
	3	---	---	2.13	4.57	4.93	1.60	2.64	1.85
	1	1.40	1.49	1.35	1.19	1.43	--	--	--
7	2	---	2.59	1.09	1.15	1.43	.89	.11	1.50
	3	---	---	.84	1.48	1.77	1.19	1.12	1.48
		(Lbs. of Fat)							
	1	5.42	3.29	3.65	8.36	3.26	--	--	--
1	2	---	.41	7.88	2.83	3.50	5.20	-1.04	1.12
	3	---	---	12.25	-6.64	13.84	1.10	3.92	2.69
	1	1.84	1.98	2.53	2.64	2.23	--	--	--
3	2	---	.61	3.79	1.55	1.14	-1.18	1.79	4.77
	3	---	---	2.01	-2.16	1.97	2.33	2.29	1.63
7	1	1.18	1.37	1.21	1.25	.73	--	--	--
	2	---	-11.50	.84	1.92	1.41	1.24	1.56	1.72
	3	---	---	1.23	-2.85	-1.56	1.34	.97	1.27

systematically to approach unity as the lactation progressed. The average intra-herd factors computed did not exhibit this characteristic except for the Holsteins. The observed fluctuation from one test day to the next could possibly be explained by the use of intra-herd regression coefficients as a measure of relationship and the relatively small numbers of records controlling their value. Regression coefficients produced by the minimum of only two records falling in a single herd, season, lactation, age group could easily change the differences between each other in their part to whole relationship from one month to the next. Consequently, if there were sufficient numbers of coefficients of this nature, the changing relationship would be reflected in the unweighted averages. When the data were scrutinized, it was found that 45-55 per cent of the intra-herd regression coefficients were the products of two observations. The number of intra-herd, -season, -lactation, -age regression coefficients for the four breeds were: 2,699 Holstein, 440 Guernsey, 208 Jersey, and 98 Brown Swiss. The number of regressions might seem sufficient to average out much of the random variation among the coefficients, but the results indicate that the variation among individual regressions was so large and the number of regressions per subclass was so small that sampling variation caused the average coefficients to fluctuate widely.

WEIGHTED EXTENSION FACTORS

When each intra-herd, -season, -lactation, -age regression coefficient was weighted by the number of records controlling its value, the trends in the average regression coefficients were smoothed out to a considerable extent. By weighting in this manner, the average regression coefficients were determined more by the regressions fitted to many records within a single herd, season, lactation, age group and the corresponding coefficients have less tendency to deviate excessively from one test day to the next. The weighted intra-herd extension factors (b) are presented in Tables 3-6.

A more accurate method of weighting would have been to weight each intra-herd, season, lactation, age coefficient by the inverse of its variance. Coefficients with the least variation would therefore be given more value in determining the average coefficient.

AVERAGE TEST DAY PRODUCTION

Means (A) and standard deviations (a) of cumulative test day production for the four breeds are also presented in Tables 3-6. Differences between herds and cows are included in these values. The average variation was less for the smaller breeds (Guernsey and Jersey) than the larger breeds (Holstein and Brown Swiss). The variation in cumulative production increased at a decreasing rate for all breeds.

Correlations (r) between cumulative test day production and the sum of the ten consecutive test days ignoring herd differences are presented in Table 7. They compare quite closely with those reported in the literature. The Holstein and Brown Swiss correlations are very closely matched and are at least .9 by the fourth month. The Guernsey and Jersey are similar to each other and although being lower in the initial stage of the lactation are .9 by the fifth month.

COMPARISON OF INTER AND INTRA-HERD EXTENSION FACTORS

The Holstein inter-herd (B - total regression ignoring herd differences)

TABLE 3. REGRESSION FACTORS FOR EXTENDING CUMULATIVE TEST DAY PRODUCTION TO A 305-DAY BASIS - Holstein*

Test Day	A	a	B	b	c
		(Lbs. of Milk)			
1	50.4	12.3	5.09	5.19	7.63
2	100.4	23.8	2.91	2.62	3.83
3	146.4	33.4	2.11	2.15	2.63
4	188.9	43.8	1.72	1.69	2.04
5	228.5	52.3	1.47	1.61	1.68
6	265.5	60.0	1.31	1.37	1.45
7	300.3	67.0	1.20	1.31	1.28
8	332.5	73.0	1.11	1.19	1.16
9	361.2	78.1	1.05	1.11	1.07
10	384.7	82.0	1.00	1.00	1.00
		(Lbs. of Fat)			
1	1.95	.59	4.04	3.32	7.12
2	3.70	1.00	2.54	2.42	3.75
3	5.30	1.36	1.96	1.89	2.62
4	6.78	1.64	1.64	1.77	2.05
5	8.17	1.96	1.44	1.51	1.70
6	9.49	2.22	1.30	1.43	1.46
7	10.73	2.46	1.19	1.28	1.29
8	11.91	2.67	1.11	1.15	1.17
9	12.98	2.86	1.05	1.08	1.07
10	13.88	3.00	1.00	1.00	1.00

* A is cumulative test day production

a is the standard deviation of cumulative test day production

B is the inter-herd regression of whole on cumulative part production

b is the intra-herd regression of whole on cumulative part production

c is the ratio of whole to cumulative part production

TABLE 4. REGRESSION FACTORS FOR EXTENDING CUMULATIVE TEST DAY PRODUCTION TO A 305-DAY BASIS - Guernsey*

Test Day	A	a	B	b	c
		(Lbs. of Milk)			
1	37.1	8.5	4.79	3.30	7.19
2	72.8	16.4	2.79	2.05	3.66
3	105.0	23.5	2.05	2.08	2.54
4	134.2	30.0	1.67	2.01	1.99
5	161.0	35.8	1.45	1.42	1.66
6	185.9	41.0	1.30	1.01	1.43
7	209.3	45.7	1.19	1.30	1.27
8	231.0	49.8	1.12	1.19	1.15
9	250.4	53.3	1.06	1.08	1.07
10	266.7	56.7	1.00	1.00	1.00
		(Lbs. of Fat)			
1	1.72	.46	3.97	2.97	7.40
2	3.32	.83	2.51	2.12	3.83
3	4.79	1.13	1.96	2.54	2.66
4	6.16	1.41	1.65	1.63	2.07
5	7.44	1.66	1.45	1.77	1.71
6	8.66	1.88	1.31	1.20	1.47
7	9.80	2.10	1.21	1.16	1.30
8	10.90	2.29	1.13	1.20	1.17
9	11.88	2.47	1.06	1.11	1.07
10	12.73	2.64	1.00	1.00	1.00

* A is cumulative test day production

a is the standard deviation of cumulative test day production

B is the inter-herd regression of whole on cumulative part production

b is the intra-herd regression of whole on cumulative part production

c is the ratio of whole to cumulative part production

TABLE 5. REGRESSION FACTORS FOR EXTENDING CUMULATIVE TEST DAY PRODUCTION TO A 305-DAY BASIS - Jersey*

Test Day	A	a	B	b	c
			(Lbs. of Milk)		
1	32.6	7.7	4.75	3.51	7.14
2	64.0	14.5	2.80	1.98	3.64
3	92.1	20.6	2.09	2.07	2.53
4	117.3	26.3	1.71	1.99	1.98
5	140.3	31.5	1.48	1.37	1.66
6	161.9	35.8	1.32	1.29	1.44
7	182.1	40.2	1.20	1.01	1.28
8	201.0	43.9	1.12	1.13	1.16
9	218.0	47.1	1.06	1.05	1.07
10	232.7	50.0	1.00	1.00	1.00
			(Lbs. of Fat)		
1	1.60	.46	4.02	3.18	7.61
2	3.14	.80	2.62	2.65	3.88
3	4.57	1.09	2.04	1.74	2.66
4	5.88	1.36	1.69	1.70	2.07
5	7.11	1.61	1.48	2.34	1.71
6	8.27	1.85	1.32	1.80	1.47
7	9.36	2.07	1.21	1.48	1.30
8	10.40	2.25	1.12	1.27	1.17
9	11.34	2.42	1.06	1.12	1.07
10	12.17	2.58	1.00	1.00	1.00

* A is cumulative test day production

a is the standard deviation of cumulative test day production

B is the inter-herd regression of whole on cumulative part production

b is the intra-herd regression of whole on cumulative part production

c is the ratio of whole to cumulative part production

TABLE 6. REGRESSION FACTORS FOR EXTENDING CUMULATIVE TEST DAY PRODUCTION TO A 305-DAY BASIS - Brown Swiss*

Test Day	A	a	B	b	c
		(Lbs. of Milk)			
1	42.9	11.2	5.11	4.03	7.79
2	85.8	22.2	2.80	3.47	3.90
3	125.9	31.7	2.07	1.66	2.65
4	163.2	40.3	1.69	1.37	2.05
5	197.8	48.1	1.46	1.60	1.69
6	230.3	55.1	1.30	1.28	1.45
7	260.5	61.3	1.19	1.29	1.28
8	288.2	66.4	1.12	1.19	1.16
9	313.1	70.9	1.06	1.07	1.07
10	334.2	75.2	1.00	1.00	1.00
		(Lbs. of Fat)			
1	1.72	.53	4.40	3.25	7.67
2	3.34	.93	2.72	2.66	3.93
3	4.86	1.29	2.08	1.76	2.71
4	6.29	1.61	1.73	1.73	2.09
5	7.63	1.91	1.50	1.52	1.73
6	8.91	2.18	1.33	1.27	1.48
7	10.12	2.43	1.22	1.19	1.30
8	11.25	2.66	1.13	.97	1.17
9	12.28	2.85	1.06	1.14	1.07
10	13.19	3.04	1.00	1.00	1.00

* A is cumulative test day production

a is the standard deviation of cumulative test day production

B is the inter-herd regression of whole on cumulative part production

b is the intra-herd regression of whole on cumulative part production

c is the ratio of whole to cumulative part production

TABLE 7. CORRELATIONS BETWEEN CUMULATIVE TEST DAY AND TOTAL PRODUCTION, AND STANDARD ERRORS OF ESTIMATE *

Test Day	Breed							
	Holstein		Brown Swiss		Guernsey		Jersey	
	r	e	r	e	r	e	r	e
				(Lbs. of Milk)				
1	.76	61.1	.76	57.1	.72	44.1	.73	38.6
2	.85	48.9	.82	48.9	.81	36.2	.81	31.5
3	.88	42.1	.87	41.2	.85	31.6	.86	26.9
4	.92	35.3	.91	35.2	.89	27.8	.89	23.7
5	.94	30.9	.93	29.8	.91	24.1	.92	20.3
6	.96	25.7	.95	24.6	.94	20.3	.94	17.0
7	.98	18.6	.97	19.1	.96	16.1	.97	13.5
8	.99	14.8	.99	13.3	.98	11.2	.98	9.6
9	1.00	7.1	1.00	7.2	.99	5.9	1.00	5.2
				(Lbs. of Fat)				
1	.77	2.6	.76	2.6	.70	2.4	.71	2.4
2	.84	2.0	.83	2.0	.79	2.0	.81	1.8
3	.89	1.6	.88	1.6	.84	1.6	.87	1.5
4	.92	1.4	.91	1.4	.88	1.4	.90	1.3
5	.94	1.2	.94	1.2	.91	1.2	.92	1.1
6	.96	1.0	.96	1.0	.94	1.0	.95	.9
7	.97	.8	.97	.7	.96	.8	.97	.7
8	.99	.5	.99	.5	.98	.5	.98	.5
9	1.00	.3	1.00	.3	.99	.3	.99	.3

* r is correlation ignoring herd, season, lactation and age differences

e is standard error of estimate ignoring herd, season, lactation and age differences and intra-herd (b) extension factors can be compared in Table 3. All intra-herd milk factors are larger than the inter-herd factors except for the second and fourth cumulative test days. The corresponding factors for fat production are larger for all test days starting with the fourth. Differences between the two sets of extension factors have not been tested. A marked similarity exists between the sets of factors, especially for milk production. Since variation due to herd differences is significant only for the first month of milk production (the probabilities are approximations, see Table 1), this suggests that herd differences are not an important factor influencing the part to whole lactational relationship. Consequently, it may not be necessary or practical to derive extension factors on an intra-herd basis to achieve maximum accuracy in extending incomplete records to a 305-day basis.

APPLICATION OF EXTENSION FACTORS

The general regression equation is more readily understandable for predicting milk and fat production from an incomplete record when in the form:

$$\hat{Y} = \bar{Y} + b(X - \bar{X})$$

Estimated production ($\hat{Y}$) is a function of the average ten month production ($\bar{Y}$) added to the product of the extension factor (b) times the difference between the incomplete production record (X) and the average production for a similar length of lactation ($\bar{X}$).

INTER-HERD FACTORS

Average cumulative monthly production ($\bar{X}$) and ten month production ($\bar{Y}$) can be obtained from Tables 3-6 by multiplying test day production (A) by 30.5, the average number of days in a month. As an illustration, imagine an estimate of 305-day milk production is desired for a Holstein cow having produced 5,000 lbs. in four months of an incomplete record. From Table 3, the average Holstein milk production for four test days (A_4) is 188.9 lbs.; multiplied by 30.5 equals 5,761.45, the average for four months production. Similarly, the ten month average production (A_{10}) is 384.7 times 30.5 or 11,733.35 lbs. of milk. Substituting these values into the predicting equation above and using the appropriate inter-herd extension factor (B) from Table 3, 11,733.35 ($\bar{Y}$) plus 1.72 (B_4) times 5,000 (X) minus 5,761.45 ($\bar{X}$) equals 10,423.66 lbs. of milk the estimated 305-day production ($\hat{Y}$).

INTRA-HERD FACTORS

The ten month lactation herd average will be available to the dairyman from his annual herd summary report, but it is not likely that he will know the average monthly cumulative production for the other nine months of the lactation for his herd. This can be estimated by using the c values presented in Tables 3-6 and his ten month lactation herd average. The c represents the ratio of ten month production to average cumulative monthly production. It may be expressed as: $c = \frac{\bar{Y}}{\bar{X}}$. Since c and $\bar{Y}$ are known quantities, $\bar{X}$ can be computed by dividing the ten month lactation herd average by the appropriate c value; i.e. $\frac{\bar{Y}}{c} = \bar{X}$. The estimate of $\bar{X}$ assumes that any herd has a similar relationship of cumulative part to whole ten month production as the breed average.

For an example of practical application, assume an estimate of 305-day milk production is desired for a Holstein cow having produced 5,000 lbs. in four months of an incomplete record in a herd with a 10,000 lb. 305-day herd average. From Table 3, 10,000 ($\bar{Y}$) divided by 2.04 (c) equals an estimated four month lactation herd average of 4,901.96 lbs. ($\bar{X}$). Substituting these values into the predicting equation and using the appropriate intra-herd extension factor (b) from Table 3, 10,000 ($\bar{Y}$) added to 1.69 (b_4) times the difference between 5,000 (X) and 4,901.96 ($\bar{X}$) equals 10,165.69 lbs. of milk the estimated 305-day production ($\hat{Y}$).

SUMMARY

Cumulative test day production for milk and fat from 11,420 Michigan D.H.I.A. - IBM records for the period 1953 to July 1957 were analysed and used to derive regression factors for extending a partial record to a 305-day basis. The data included 8,993 Holstein, 1,457 Guernsey, 651 Jersey , and 319 Brown Swiss records.

Linear regression coefficients were used as the variable measuring the relationship between cumulative production on test day and the sum of the first ten test days. The regression coefficients were calculated within each herd, season, lactation, and age group. Analyses of variance and visual inspection of the milk and fat regression coefficients for the first, third, and seventh cumulative months of the Holstein data were used to determine the importance of season, lactation, and age on the part to whole relationship. From these techniques, season was not considered an important influence for the present study. Visual inspection of the mean age coefficients within lactation numbers failed to reveal where the groups should be separated or combined; consequently, the data were grouped as a whole within each breed.

Intra-herd extension factors were computed as the weighted average regression coefficient within herds, seasons, lactations, and ages for each of the nine cumulative test days. Each regression coefficient included in the average was weighted by the number of records determining its value. For comparison, inter-herd extension factors were also derived disregarding herd, season, lactation, and age effects. Although the inter and intra-herd factors were not tested for significance, the marked

similarity between the Holstein factors suggests that herd has little or no influence on the part to whole relationship.

Correlations between cumulative part and 305-day production averaged over all herds, seasons, lactations, and ages were not less than .7 for the first month, increased steadily as the lactation progressed; and were .9 by the fifth test day for all breeds. This supports the literature that suggests production records of only one or two months are valuable guides to what a cow will produce in that lactation. Furthermore, extension of such records to a 305-day basis would be useful tools in early culling and progeny testing.

REFERENCES

(1) Becker, R. B., and Arnold, P. T. Dix. Influence of Season and Advancing Lactation on Butterfat Content of Jersey Milk. J. Dairy Sci., 18: 389. 1935.

(2) Cannon, C. Y., Frye, J. B., Jr., and Sims, J. A. Predicting 305-Day Yields from Short-Time Records. J. Dairy Sci., 25: 991. 1942.

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

(4) Gifford, W. The Relative Accuracy of Various Portions of the Lactation as Indicators of the Permanent Productivity of Cows. Unpublished Ph.D. Thesis. Iowa State College Library, Ames. 1939.

(5) Harvey, W. R. Unpublished data from Ayrshire Herd Test records. 1956.

(6) Kendrick, J. F. Unpublished data from Ayrshire Herd Test records. 1955.

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

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

(9) __________, McGilliard, L. D., and Ralston, N. P. Relations Between Monthly Test-Day Milk Production of Holstein-Friesian Cows. J. Dairy Sci., 39: 932. 1956.

(10) __________, McGilliard, L. D., and Ralston, N. P. Relations Between Test-Day Milk Production of Holstein Cows. J. Dairy Sci. 1959. (In press).

(11) Michigan D.H.I.A. - IBM Annual Herd Summary. Mimeo. 1955.

(12) __________ Annual Herd Summary. Mimeo. 1956.

(13) __________ Annual Herd Summary. Mimeo. 1957.

(14) Rendel, J. M., Robertson, A., Asker, A. A., Khiskin, S. S. and Ragab, M. T. The Inheritance of Milk Production Characteristics. J. Agr. Sci., 48: 426. 1957.

(15) Voelker, H. H. Use of Extended Incomplete Lactation Records. J. Dairy Sci., 40: 631. 1957.

(16) Wylie, C. E. Factors Influencing Two- Day Official Butterfat Tests of Cows. J. Dairy Sci., 6: 294. 1923.