LON D. McGILLIARD

MEASURING YEARLY ENVIRONMENTAL DIFFERENCES WITHIN A HOLSTEIN HERD

Thesis for the Degree of M. S.

MICHIGAN STATE COLLEGE

Lon D. McGilliard

1947

This is to certify that the

thesis entitled

"Measuring Yearly
Environmental Differences
Within A Holstein Herd"

presented by

Lon D. McGilliard

has been accepted towards fulfillment
of the requirements for

M. S. degree in Dairy

Earl Weaver
Major professor

Date December 18, 1947

M-795

MEASURING YEARLY ENVIRONMENTAL DIFFERENCES
WITHIN A HOLSTEIN HERD

by

Lon D. McGilliard

1947

MEASURING YEARLY ENVIRONMENTAL DIFFERENCES

WITHIN A HOLSTEIN HERD

By

Lon D. McGilliard

A THESIS

Submitted to the School of Graduate Studies of Michigan
State College of Agriculture and Applied Science
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE

Department of Dairy

1947

THESIS

ACKNOWLEDGMENTS

For their invaluable assistance and suggestions in the development of this project and in the preparation of this paper, the author wishes to express his sincere appreciation to Dr. Earl Weaver, Head of the Department of Dairy, and to Dr. Ronald H. Nelson, Associate Professor of Animal Husbandry.

Gratitude is also expressed to Dr. Leo Katz, Associate Professor of Mathematics, for his aid in the statistical treatment of the problem.

TABLE OF CONTENTS

INTRODUCTION

The problem of determining the extent of distortion that varying environmental effects produce on the true hereditary picture in milk production of dairy cattle is extremely useful yet particularly difficult.

The usefulness of a successful study of this problem is almost unlimited. Mistaking the effects of environment for the effect of heredity is generally the largest barrier to the breeder's rapid attainment of his goal. If variable environmental effects upon milk production could be equalized, the breeder would then have a clear picture of the value of his breeding program, one unobscured by outside influences. He would have a reliable estimate of his animals to aid in selection and mating toward the improvement of his herd. By the same theory he would be able to place a more reliable index upon his herd sires, one which would be of greater use and reliability. The whole course of a breeding program would be clarified were it not for the confusing effects produced by non-hereditary factors.

In dairy cattle the general criterion of productive ability is the amount of milk or the amount of butterfat which a cow is able to produce in a given period of time. This quantity is used not only to represent the potentiality of the individual cow, but it is also incorporated in the index by which the value of her sire is judged.

The amount of milk and fat a cow can produce in a lactation period depends upon a multitude of factors. These factors may be divided into two general classes, hereditary factors and environmental factors. The hereditary factors include all genetic effects while all else not inherent is grouped under environment. So that the issue not be confused, the assumption is made that non-additive environmental interactions are non-existent.

Various estimates of the heritability of fat production have been made. Among these there has been some variation but in general the estimates are similar. The usual figure is below 40 per cent for the heritability of fat production. Lush (2) mentions that intrasire regression of daughter on dam have yielded values of around 0.15 to 0.30 for heritability of differences in fat production between cows in the same herd where each cow is represented by one record. The remaining variance of fat production is then grouped under environmental effects. It is easy to imagine how the effects of environment could be confused with heredity and complicate breeding and selection operations greatly.

In general the observed variance may be described by the equation, $\sigma_H^2 + \sigma_E^2 = \sigma_O^2$ where σ_H^2 represents the variance due to heredity and σ_E^2 is variance due to environment. Obviously a change in one of the first members produces a corresponding change in the opposite direction in the other. Thus by a decrease of σ_E^2 an increase in σ_H^2 results. Conversely, if a greater proportion of observed variance is due to environmental variations, variance due to hereditary differences will be lessened. In order to increase the value of σ_H^2, it is necessary that σ_E^2 be reduced to a minimum. If variance due to differences in heredity can be made to constitute the entire observed variance, breeding and selecting processes will be much simplified. The primary purpose of this study is to attempt to determine a method whereby environmental variance can be reduced to a minimum to produce a corresponding clarification of hereditary variance.

There are two methods possible for reducing the σ_E^2 , the physical control of environment and statistical correction for environmental effects. The first of these would require testing all animals under standard conditions to reduce σ_E^2 to zero. Obviously complete physical control of environment is an impossibility. Even partial control may be very impractical.

Statistical control seems a little more plausible under present conditions. Such control would consist of using correction factors to allow for unusual conditions varying from standard or to equalize the conditions under which two animals may have been. The breeder performs this operation to some extent mentally when he knows and allows for the fact that one animal performed under more adverse conditions than another or when he recognizes that his feed was inferior quality one year, or that one cow had a short dry period, or an abortion one lactation. The range of corrections would run from this type up to very complicated computations involving too much labor for the accuracy involved. It is not unreasonable to believe that some of these corrections might actually induce added error, but ordinarily they should be usable if they are more often correct than wrong.

Some correction factors in the intermediate range are already in use in dairy work. Examples of these are the corrections for lactation length, for the frequency of milking, and for age, the latter being the mature equivalent correction. Another example might be the separate production classes for animals meeting certain calving requirements. These are all corrections to aid in equalizing environmental effects, and some reliability is attached to them. It is true they are composed from averages and may be radically incorrect for some individuals, but these errors will tend to cancel each other when numbers of individuals are involved.

An attempt was made by Johannson (1) to correct for the total environmental differences between years in a dairy herd by a relatively simple method which will be described fully later in the paper. This method was again presented along with a complicated statistical method for which there was a test of significance by Nelson (3). The results of the two methods agreed; therefore, the simplier was selected as the desirable method.

This paper will present several different methods which might be used to remove the environmental differences between years by the calculation of correction factors to be applied to herd averages. When corrected, the herd average should present a picture of the decline or improvement in hereditary merit of the herd.

THEORY OF ENVIRONMENTAL DETERMINATION

The principle involved in the determination of environmental effects is that any individual possesses an inherent capacity for performance. Whether each individual reaches the full extent of its productive capacity is limited by all factors other than inherent, in other words environmental. The individual has the ability to produce at a certain level under optimum conditions. Any variation from optimum involves a reduction in performance below maximum.

Under this theory if we consider the production of a cow during one lactation and assume that to be her capacity for production, then any other of her lactations under the same conditions should be equal. Any variation between them should be due to one or more of the effects which shall be classified as environmental. If these differences between lactations take the form of corrections to be applied to the herd averages in their respective years, then the corrected herd averages will be equal as far as environmental differences are concerned; that is, all environmental differences will be removed. Remaining differences are those due to genetic merit. Any fluctuation of average would demonstrate hereditary improvement or decline in the herd. It is not impossible to believe that this might be carried further and this fluctuation pinned to specific sires perhaps as an aid in formulation of sire indices.

SOURCE OF DATA

The data for this problem were drawn from the records of the Michigan State College Dairy Department. Since the early nineteen hundreds a Holstein herd has been maintained as a part of the college dairy herd for which fairly complete records have been kept, especially in regard to production. Although five dairy breeds are maintained by the department, the Holsteins were selected for this study because of the larger number of animals.

The author had occasion recently to combine all records of the Holsteins in the college herd; these included breeding, disposal, calving, and production records. The production records for this work were drawn from this information.

At first it was planned to conduct the study from 1930 to 1946. The year, 1930, was selected as a starting point since it was prior to this year that the dairy herd became seriously affected with contagious abortion. During this and the period immediately preceding 1930 all infected animals were removed from the herd. In 1929 and 1930 the nonreactors were moved across the river to a new barn. All reactors remained behind and were eliminated. Only a very few of the original herd were moved to the new barn. Several purchases were made to replenish the herd, these animals entering the herd in 1930.

The first plan was to include all animals in the herd in 1930 and to drop back before 1930 to pick up any records which the transfer cows had made prior to then. This plan was rejected when it was observed that there were too few transfer cows, and also because of the complete change of conditions with the move to the new barn. Consequently, 1930 was chosen as the first year.

After all of the records were collected, it became apparent that the production of the purchased cows was considerably higher in 1930 than the production of the transfer group and was also higher than any subsequent lactations of the purchased cows. An investigation as to the cause revealed that the better cows were milked four times a day in1930 while some were milked partly 4x and partly 3x; others were milked 2x and 3x entirely. However, there seemed to be no record of the exact number of milkings of each frequency for each cow. Also some cows were milked by hand and some, by machine. Starting with 1931 and continuing to now all cows were milked three times a day. In view of this, it was decided to use 1931 as the starting year; consequently, records of 1931 to and including 1946 were used.

During this entire period all cows in the herd were milked three times a day, and the milk weights were recorded for each milking. Monthly fat tests were made, and butterfat production was computed.

Over this span of years the herd has been maintained largely by replacements bred and raised by the college. Very few animals have been introduced from outside sources. There presumably has not been too much selection and culling although as will be shown later, closer culling was practiced than was supposed at first. General conditions were thought to have varied somewhat but not extremely. The same herdsman was in charge for the entire period.

In the compilation of records, it was observed that most lactations were longer than 305 days, while some were shorter. All records of less than 270 days were eliminated although there may be some question as to whether they should be as they may have been shortened by some environmental effect. The best plan would be to deal with them seperately as they are not comparable with the 305 day records, and they may induce considerable error in calculations if included. All lactations of more than 305 days were reduced to 305 days by interpolating in the month in which the 305th day occurred and totaling with the preceding monthly amounts.

Each record was included in the year in which that lactation was initiated, the year being the calendar year. Actually some of the lactations recorded as a part of one year may have taken place mostly in the following year; for instance a lactation beginning in December of 1931 was recorded as 1931, but nine months of the production really occurred in 1932. It is unavoidable that some overlapping of records should occur with a system such as this where cows are freshening every month of the year, and, unfortunately, this overlap of records masks the definite demarcation of yearly effects.

Since, as it was mentioned previously, it is necessary that all records be on an equivalent basis to remove the age effect, the records were all converted to "mature equivalents." The factors used for conversion were the same as those furnished by the IHIA for field use. The smallest division was three months; further accuracy was unnecessary since the entire problem is an approximation. That conversion left the records in working form as 3x, 305-day, M.E. records. There is always a question as to the reliability of mature equivalent records, which argument belongs not to this discussion. It must be assumed that mature equivalent records are sound and reliable for this study is based upon that assumption which is the foundation of the theory of environmental correction.

PLAN I - PAIRED YEAR COMPARISONS

This is the system mentioned earlier as having been advanced by Johannson (1) and tested by Nelson (3). Nelson found similar results were obtained by this as by a more extensive plan and concluded that the simplicity of the method made it more desirable than the longer method.

This plan is developed under the basic assumption that all of any cow's lactation records of butterfat production if equal in length and frequency of milking, and if converted to remove age differences, will be equal under identical environmental conditions. In other words if a given cow produces 600 pounds of butterfat on 3x, 305-day, M.E. in 1931, her record in 1932 on this same basis and under identical conditions should be 600 pounds. Likewise, her 1933, 1934, and 1935 records should be the same. The inherent ability of the individual is determined by the genetic makeup of that individual and is not variable within the individual. If, however, there is some variation between different lactations of the same cow, some factors or conditions other than inherent must have been present to alter the production. These are grouped under the title of environment, and it is these differences which are to be determined.

The actual operation of this plan is very simple. It consists of grouping the M.E. records of cows with lactations in two successive years in their respective years and calculating the mean difference between the production of the two years. The process would be:

$$\frac{\sum (x_{i(2)} - x_{i(1)})}{N}$$

where $X_{1(1)}$ equals the production of a cow in the first of two succeeding lactations; $X_{1(2)}$ equals the production of the same cow in the succeeding year; and N is the number of comparisons. For example, from the theoretical data in Figure 1, the differences are calculated in Figure 2. The result indicates that the environment of year, 1932, decreased the production of the same cows an average of 100 pounds under 1931 while 1933 was 100 pounds better than 1932.

	1931	1932	1933
Cow A	600	500	600
Cow B	500	400	
Cow C		600	700

Figure I. THEORETICAL BUTTERFAT RECORDS

	1931	1932	1932	1933
Cow A	600	500	500	600
Cow B	500	400		
Cow C			600	700
Total	1100	900	1100	1300
$\frac{\sum (X_{1(1)} - X_{1(2)})}{N}$	$\frac{-200}{2}$	$= -100$	$+\frac{200}{2}$	$= +100$

FIGURE II. COMPUTATION OF ENVIRONMENTAL DIFFERENCE FROM DATA IN FIG. I.

This part of the procedure agrees with the methods set forth by Nelson and Johannson. Some difference lies in the use of the factor as a correction to be applied to the herd average. Johannson developed a multiplicative factor to be used in the correction of the herd averages. Nelson's method was to select a base year, to determine a relative deviation of each year from the base by the addition of deviations from year to year, and to use these deviations as corrections for the herd averages by reversing the signs. This is the procedure used in this study. The first year, 1931, was selected as the base year only because the use of the initial year simplifies the process in that the additive correction proceeds only in one direction. However, it might be desirable to use the highest plus year for the base with all corrections figured in comparison with it, but the same relative result is obtained either way.

The results of this first plan are shown in Table I. The column of sums of differences, (6) is a cumulative difference for each year by which each year is placed on the same basis, in this case by comparison with 1931. The herd averages shown are the average records of only the cows completing the full lactation started in each particular year. They include a few individuals who completed only one lactation and were not used in the evironmental difference calculations. The corrected average column (8) contains the herd average for each year after the calculated environmental difference has been applied to the herd average.

It is obvious that a very distinct trend toward the negative exists in the data presented. Only four positive differences show, and those are not great enough to counteract the negative trend. Column 6 of the table demonstrates that the swing to the left is extreme and much out

TABLE I. PLAN I-A. PAIRED YEAR COMPARISONS

1 Year	2 No. of Cows	3 Preceding Years Av.	4 Present Years Av.	5 Difference (3) - (4)	6 Sum of Diff. (Cum. 5)	7 Actual Herd Av.	8 Corrected Av. (7) - (6)
1931					0	511	511
32	11	560	481	- 79	- 79	507	586
33	8	540	533	- 7	- 86	570	656
34	12	564	559	- 5	- 91	561	652
35	5	548	536	- 12	-103	566	669
36	4	579	470	-109	-212	506	718
37	9	501	524	+ 23	-189	528	717
38	6	565	558	- 7	-196	529	725
39	5	595	622	+ 27	-179	567	745
40	8	576	564	- 12	-191	565	756
41	8	553	509	- 44	-235	527	762
42	5	546	551	+ 5	-230	565	795
43	9	565	512	- 53	-283	523	806
44	10	513	508	- 5	-288	530	818
45	10	533	543	+ 10	-278	523	801
46	4	555	540	- 15	-293	643	936

Total No. of Records - 114

of proportion. This is further emphasized in column 8. It is entirely unreasonable even to presume that the merit of the herd increased from 511 pounds of butterfat in 1931 to 936 pounds in 1946 as indicated or that the environment was so extremely limiting in the last years. Some factor seems to be working in the data to produce this severe irregularity.

When this negative swing was noted, the first possible explanation to come to mind was that the mature equivalent factors were unsuitable for this population. Since the comparison between each two years moves progressively forward, with each comparison young individuals enter into the first year average while the next year they are one year older. That might indicate that the conversion factors were too high for the younger ages since such a situation would produce the negative trend. This is one of the criticisms of mature equivalent factors usually offered.

An effort was then made to check the suitability of the M.E. factors used for this population by computing factors from the records of the animals used in this study. The results are shown in Table 2.

TABLE II. COMPARISON OF MATURE EQUIVALENT CONVERSION FACTORS

Age	2-0	2-6	3-0	3-6	4-0	4-6	5-0	5-6	6-0	6-6	7-0
No. of Comparisons	9	5	7	9	10	9	10	7	10	10	10
Computed M.E. Factors	1.448	1.241	1.351	1.174	1.210	1.139	1.039	1.095	1.000	1.000	1.000
DHIA M.E. Factors	1.377	1.275	1.203	1.131	1.077	1.035	1.017	1.006	1.000	1.000	1.000

In spite of the small numbers used, these computed factors seem to be fairly comparable with the ones used; if anything, they do not make a great enough correction for the younger individuals. This evidence lessened the possibility of unsuitable mature equivalent factors influencing the results markedly.

The second possibility suggested was that selection may have been more rigid than was supposed, thus causing the observed trend in the results. The theory behind this is when one record is used as the indication of the level of an individual's producing ability, the next record can be expected to regress toward the population average. If selection has been practiced to any extent in the herd, the selected individuals will be above the herd average; consequently, their subsequent lactations will tend to regress toward the average of the herd.

A check on the selection employed can be made quickly by a comparison of the records of cows retained in the herd with those removed from the herd. The herdsman and others connected with the management of the herd insist that there has been practically no selection, that the problem of maintaining sufficient numbers in the herd has been difficult. Table 3 illustrates that the average production of the cows retained in the herd as compared with the herd average for each year is consistantly higher. In some of the years the better cows passed out of the picture, pulling the retained cow average down, but especially in 1931, 1932, and 1938 was there a great difference between cows retained and cows removed.

This evidence makes selection a likely factor in producing the bias in the correction factors. An error introduced in the early years is carried through to the last year plus any errors introduced in the intervening years; consequently, the end year accumulates all of the errors. The next question is what can be done to eliminate some of this error.

TABLE III. COMPARISON OF AVERAGE BUTTERFAT PRODUCTION OF COWS RETAINED IN THE HERD WITH THE HERD AVERAGE

	1931	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
No. Cows	22	17	17	13	10	12	15	9	13	14	10	12	13	15	13	9
Herd Av.	511	507	570	561	566	506	528	529	567	565	527	565	523	530	523	643
No. Retained	13	9	14	10	4	9	8	5	8	10	5	9	10	12	4	-
Av.	547	541	569	558	579	501	545	595	576	552	546	565	515	539	555	-
No. Removed	9	8	3	3	6	3	7	4	5	4	5	3	2	3	9	-
Av.	458	469	574	575	558	521	509	449	553	598	507	566	569	494	509	-
Retained Av. Minus Herd Average	+36	+34	- 1	- 3	+13	- 5	+17	+66	+ 9	-13	+19	0	- 8	+ 9	+22	-

According to Lush (2) the repeatability of yearly fat production considering only cows in the same herd is somewhere between 1/3 and 1/2, and he suggests the use of the general figure, 2/5, although a higher estimate might be justified in herds with standardized management and relatively complete environmental corrections. This repeatability (r) or correlation coefficient between records made by the same cow is the part of the total variance between corrected records which is due to permanent differences between cows.

Where the repeatability of production is low, the use of one record as the indication of the merit of the cow is not reliable. To estimate the probable producing ability of a cow, Lush offers the equation:

$$\text{Probable producing ability} = \frac{NR}{1 - r + nr} \times \text{(cow's record)} + \frac{1 - r}{1 - r + nr} \times \text{(herd average)}.$$

Using one record and the repeatability factor, 2/5/, the equation becomes: Cow's ability = 2/5 (her average) + 3/5 (herd average). This equation seems useful in this case where all known environmental effects have been corrected, and there is still evidence of low repeatability of production.

It was decided to try applying this regression factor to the study to determine if it would bring the corrected herd averages within a reasonable range. The factor was applied to the first year of each comparison, the reasoning being that on the basis of the one record and a repeatability of production of 2/5, the second record could be predicted at 3/5 closer to the herd average than the first. Any remaining difference or the difference between the predicted and the observed amount would then be due to temporary environmental effects. The results of this study are shown in Table 4.

This operation reduced the extreme negative trend shown in the results of the first table and produced several more plus differences between years, but in comparison with the conditions of 1931, all of the years are still poorer. It is very difficult to believe that in 1946 conditions were so inferior to those of 1931 as to reduce production 150 pounds per cow. Of course, the number of cows in the study is very small and the discrepancy may lie in that since, as will be shown later, there is more variance between cows within years than between years.

Johannson (1) and Nelson (3) obtained fair results with this method in their studies. There was no apparent selection practiced in the herds they studied, and no regression factor was used. The method perhaps has possibilities; it is a simple and short means of estimating environmental effects from year to year, but lists as its disadvantages that it does not use all records and the lack of a test of significance to determine if there are actual environmental differences for which a correction need be made. Where selection is practiced as occurs in nearly all herds except

TABLE IV. PLAN I-B. REGRESSION CORRECTED PAIRED YEAR COMPARISONS

1 Year	2 No. of Cows	3 Preceding Year's Av. (Regr.Corr	4 Present Year's Av.	5 Difference (3) - (4)	6 Sum of Diff. (Cum.5	7 Actual Herd Av.	8 Corrected Av.
1931	-	-	-	-	0	511	511
32	11	531	481	- 50	- 50	507	557
33	8	520	533	+ 13	- 37	570	607
34	12	568	559	- 9	- 46	561	607
35	5	556	536	- 20	- 66	566	632
36	4	571	470	-101	-167	506	673
37	9	504	524	+ 20	-147	528	675
38	6	543	558	+ 15	-132	529	661
39	5	555	622	+ 67	- 65	567	632
40	8	571	564	- 7	- 72	565	637
41	8	560	509	- 51	-123	527	650
42	5	535	551	+ 16	-107	565	672
43	9	565	512	- 53	-160	523	683
44	10	514	508	- 6	-166	530	696
45	10	531	543	+ 12	-154	523	677
46	4	536	540	+ 4	-150	643	793

Total No. of Records - 114

experimental, another variable, regression, enters in and may induce more error. The method should provide a very rough approximation of the year to year environmental differences within a herd, but the reliability in this case is questiomable.

PLAN II - DIRECT COMPARISONS

One disadvantage of the first plan was that it did not make full use of all the records of a cow. For instance a cow produced in two consecutive years, skipped the third year for some reason, and then had a lactation in the fourth. The lactation in the fourth year had to be discarded because the only comparison usable was between consecutive years.

Another feature of Plan I which seemed as if it might be improved was that by making comparisons of only two consecutive years and then basing all corrections on the first year, the final year's correction was an accumulation of any errors made in previous corrections.

The attempt to improve these defects yielded Plan II. The solution seemed to be in developing a method of trying all of an individual's lactations to her first. If all of a cow's lactations were compared directly to the first year, that would correct both of the above situations. With a direct comparison to the base year, records following "blank " years would then be usable making all records usable. Also direct comparisons would destroy the cumulative error in the final year.

Direct comparison was the answer to the problem except that no cow produced the full fifteen years. One group produced for a certain period of years, dropped out, and a new group produced from there. This was progressive, of course, with a few dropping and a few being added each year. Consequently, there was no direct comparison of 1946 to 1931; in fact, the only direct comparisons with 1931 were the years up to 1937. The year, 1932, also could be compared with all years to 1937. Year 1933 was comparable with all years up to 1939. The question then arose as to how these direct comparisons could be linked together and related back to 1931.

The first attempt at linking these direct comparisons is shown in Table 5. As before, 1931 was used as the base year. Line 1 shows the comparisons of the average production in 1931 with the average production of the same cows in each subsequent year. The encircled numbers indicate the number of cows from the original year who had records in each particular year. In the line labded "differences" the average of all differences for the year is taken and used as the environmental difference between that year and 1931. Of course the only comparison available for 1932 is the direct comparison with 1931 which is - 79. This figure is used as the value for 1932.

Line 2 consists of direct comparisons with 1932. Since the yearly averages are now compared with 1932, 1932 becomes the base in this case. As has already been determined, the representative value for this year is - 79; therefore, the difference between 1932 and 1933 is applied to this base to make the value indirectly the difference between 1933 and 1931. Likewise the actual difference between each succeeding year and 1932 is added to - 79, and the sum is recorded in its respective column. For example the actual difference between 1932 and 1933 is - 7, applied to the base value of 1932, it gives the value, - 86, which is recorded under 1933 to indicate by this comparison 1933 was 86 pounds worse environmentally than 1931. The actual difference between 1932 and 1934 is + 57, applied to the base - 79, it gives a value of - 22. This process carries the comparison back to 1931; so each figure shown is the comparison with 1931.

TABLE V. PLAN II-A - WEIGHTED AVERAGES OF DIRECT COMPARISONS

	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946
1	x	(11) -79	(8) + 5	(6) +27	(4) -38	(3) +33	(2) +115									
2		x	(8) -86	(6) -22	(4) - 4	(3) -88	(1) -185									
3			x	(12) -46	(6) -15	(6) -156	(5) -150		(1) +108							
4				x	(5) -34	(6) -82	(5) - 98		(1) +186							
5					x	(4) -132	(2) -253	(1) -203								
6						x	(9) - 73	(4) - 61	(4) - 31	(2) - 7						
7							x	(6) -105	(6) - 89	(3) -129	(1) - 48	(2) - 71	(2) -123	(2) -147		(1) + 68
8								x	(5) - 70	(4) - 92	(2) -153	(2) -120	(3) -173	(2) -196		(1) - 40
9									x	(8) - 54	(6) -116	(5) - 51	(4) - 91	(4) -101	(1) - 56	(2) - 22
10										x	(8) -115	(6) - 68	(4) -111	(4) -121	(1) -132	(2) + 22
11											x	(5) -111	(4) -128	(4) -152	(2) -171	(3) - 99
12												x	(9) -132	(9) -133	(4) - 72	(4) - 24
13													x	(10) -129	(6) -135	(4) - 31
14														x	(10) -123	(6) - 44
15															x	(4) -134
16																x
Weighted Differences	Base	(11) -79	(16) -41	(24) -22	(19) -23	(22) -96	(24) -98	(11) - 98	(17) -42	(17) -71	(17) -116	(20) -80	(25) -124	(35) -133	(24) -119	(27) -48

Since all of the differences are alike in that they are the result of comparison with the first year, they may be combined in an average to produce a representative value for each year. As - 79 was the value of 1932, so the eight comparisons of + 5 may be averaged with the eight of - 86 to obtain a base value of - 41 for 1933. This figure is used as the base fof the direct comparisons with 1933 given in line 3. In the same manner a direct comparison between 1933 and 1934 yields - 5 which is applied to the 1933 base, - 41, to give - 46. The final value for 1934 in relation to 1931 is the weighted average of the direct comparison with 1931, the indirect comparison through 1932, and the indirect comparison through 1933. This same general procedure is used right on to the last year.

The values in the difference line are then the environmental differences between each year and 1931. They are, as explained before, averages of all the comparisons between that year and every preceding year, each referred back to the base year, 1931. If a comparison between any two particular years is desired, it may be obtained by taking the difference between the values for those years.

A few questions arose in working with this procedure. One was whether a straight average or a weighted average would be more desirable in calculating the base values for the various years. Ordinarily a weighted average would be the only method, but in this situation there was a question of whether a greater number of records was more valid than a direct comparison. For example, in 1937 the direct comparisons with 1931 totaled 2. The comparisons of 1937 with 1936 and indirectly to 1931 totaled 9. In a weighting system based on frequencies the latter is weighted 4 1/2 times the first. Should the 9 indirect comparisons be weighted 4 1/2 times 2 direct comparisons where a direct comparison is certainly more valid than an indirect? It was finally decided that the average weighted as to frequency, in the absence of a method of determining just what weight each value should receive, would be more nearly correct.

A second question was whether the regression factor was applicable in this procedure. The first method was a straight comparison between each two years, but in this case all records were used and averages of the comparisons were used. It might seem that in the use of averages the regression value might not apply, but on the other hand every comparison is between only two lactations, and if the first varied from the herd average, the one following could be expected to be nearer the herd average. In this case correction for regression might be in order.

Table 6 shows the results obtained when correction for regression was applied. The regression factor was used on the first year of each comparison to predict the probable production for the following years, and differences of the actual records from this probable value were taken. The differences show the same general trend, as might be expected, but they are smaller.

The third consideration was that this system utilized the same records in several different comparisons, these comparisons being averaged to provide a value for each year. For example, 6 of the 8 records used in comparisons of 1933-1931 are used again to compare 1933 with 1932; hence these 6 records in 1933 are compared with the base year 6 times directly and 6 times indirectly to constitute 12 of the 16 comparisons making up

TABLE VI. PLAN II-C - REGRESSION CORRECTED DIRECT COMPARISONS

	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946
1	x	(11) -50	(8) +31	(6) +44	(4) +23	(3) +47	(2) +86									
2		x	(8) -37	(6) +19	(4) +23	(3) -37	(1) -177									
3			x	(12) -12	(6) +13	(6) -108	(5) -104		(1) +46							
4				x	(5) -10	(6) -58	(5) -73		(1) +88							
5					x	(4) -90	(2) -226	(1) -152								
6						x	(9) -40	(4) -26	(4) - 1	(2) +33						
7							x	(6) -56	(6) -34	(3) -51	(1) +16	(2) -18	(2) -70	(2) -94		(1) +132
8								x	(5) +13	(4) -11	(2) -57	(2) -21	(2) -73	(2) -97		(1) +105
9									x	(8) - 8	(6) -64	(5) - 9	(4) -50	(4) -60	(1) -44	(2) +53
10										x	(8) -62	(6) -12	(4) -55	(4) -65	(1) -71	(2) +76
11											x	(5) -42	(4) -60	(4) -84	(2) -103	(3) -12
12												x	(9) -73	(9) -74	(4) -36	(4) +55
13													x	(10) -75	(6) -83	(4) +51
14														x	(10) -63	(6) +18
15															x	(4) -62
16																x
Environmental Differences																
	Base	(11) -50	(16) - 3	(24) +10	(19) +11	(22) -60	(24) -71	(11) -54	(17) - 1	(17) -11	(27) -58	(20) -20	(25) -64	(35) -75	(24) -66	(27) +28

the average value for 1933. Likewise 4 of the 6 records in 1934 compared with 1931 are used in comparison with 1932 and all of them are compared with 1933 causing comparisons of those six records to appear 16 times in the 24 composing the yearly average.

In the analysis of data it is desirable to utilize all information to the fullest extent, but ordinarily there should be no repetition. This repeated use of records may be justifiable in the manner in which they were employed here, but an investigation was made to determine the effect of eliminating this repeated use of records in various comparisons. Table 7 gives the results of this work. The only lactations used in the comparison with each base year were records of cows initiating their production in the herd that particular year. The records of cows producing in the herd appear only in comparison with their first year of lactation in the herd. Their records are used in no other direct comparisons except with their first year of lactation; hence they appear only once in a yearly average. For example, cows entering the herd in 1932 have comparisons only on line 2 in which each subsequent record of those cows is compared with the 1932 record to compose differences entering into each respective yearly average. Those entering in 1933 show only in comparison with 1933. Each individual record appears only once in comparison. These results again show a trend which is somewhat similar to the previous methods.

Table 8 is the final result of this method. It embodies the utilization of all records corrected for regression with a reduction of cumulative error in the last year, and without repetition of records in comparisons. As for the first method, there is no test of significance for this method.

TABLE VII. PLAN II-D - NON-DUPLICATING DIRECT COMPARISONS

	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946
1	x	(11) -79	(8) + 5	(6) +27	(4) -38	(3) +33	(2) +115									
2		x	(2) -308	(2) -71	(1) -130	(1) - 7										
3			(5) x	(2) -98	(3) -48	(3) -100	(3) -188		(1) +91							
4				x												
5					x	(1) -229	(1) -259	(1) -234								
6						x	(4) +10	(3) - 7	(3) -24	(2) +34		(1) +102	(1) + 19	(1) - 5		
7							x	(2) -14	(2) +20	(1) -101	(1) - 5	(1) - 63	(1) - 84	(1) -109		(1) +111
8								x	(1) +116	(1) +62	(1) -99					
9									x	(4) +32	(4) -11	(3) + 19	(2) - 8	(2) - 4	(1) + 11	(1) + 35
10										x	(2) -39	(1) +40				
11											x	(1) + 6	(1) + 18	(1) -61	(1) -93	(1) -72
12												x	(4) - 60	(4) -33	(2) +109	(1) +175
13													x	(1) -71	(2) -129	
14														x	(5) -47	(2) -63
15															x	
16																x
Environmental Difference																
Base		(11) -79	(10) -58	(13) -36	(7) -54	(8) -55	(10) -55	(6) -47	(7) +25	(8) +20	(8) -28	(7) +20	(9) -34	(10) -39	(11) -32	(6) +21

TABLE VIII. PLAN II-E - REGRESSION CORRECTED, NON-DUPLICATING DIRECT COMPARISONS

	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946
1	x	(11) -50	(8) +31	(6) +44	(4) +23	(3) +47	(2) +86									
2		x	(2) -152	(2) +32	(1) -78	(1) +45										
3			x	(5) -49	(2) + 9	(3) -105	(3) -157		(1) +43							
4				x												
5					x	(1) -153	(1) -183	(1) -158								
6						x	(4) +58	(3) +29	(3) +12	(2) +58		(1) +98	(1) +10	(1) -14		
7							x	(2) +32	(2) +66	(1) -34	(1) +62	(1) + 4	(1) -17	(1) -42		(1) +178
8								x	(1) +163	(1) +109	(1) -52					
9									x	(4) +34	(4) - 9	(3) +31	(2) - 6	(2) - 2	(1) +10	(1) +34
10										x	(2) -21	(1) +50				
11											x	(1) +33	(1) +45	(1) -34	(1) -66	(1) -45
12												x	(4) -39	(4) -12	(2) +82	(1) +261
13													x	(1) -53	(2) -107	
14														x	(5) - 2	(2) - 6
15															x	
16																x
Environmental Difference		(11) -50	(10) - 6	(13) + 6	(7) + 5	(8) -35	(10) -25	(6) - 1	(7) +53	(8) +41	(8) - 8	(7) +39	(9) -14	(10) -20	(11) -11	(6) +69

PLAN III - ADJUSTED MEAN DIFFERENCES

The last method is a relatively complicated system requiring considerable calculation, but carrying a test of significance. It is offered as a possibility for more exacting environmental determinations and as a check on the two preceding plans.

Nelson (3) found the least squares method of fitting constants as shown by Yates (5) to be applicable to similar data to these in environmental difference determinations. The method was satisfactory except for the extremely lengthy calculations.

Since these data were similar in being adaptable to a two way table of classes containing disproportionate numbers, it was decided to use the least squares method of fitting constants to analyze these data. However, an alternate method was offered which seemed much simpler and was adopted for this trial.

The method used is that described by Patterson (4) as a method of adjusting which when applied to data with unequal subclass numbers, makes it possible to obtain a sum of squares for each source of variance that is free of the influence of the other effect. Likewise adjusted border means may be calculated by the method of adjusting, such means being devoid of the confounding effect produced by disproportionate subclass numbers. A complete discussion of the method may be had by referring to the original article. This paper will be confined principally to a description of the actual process involved.

The problem is to adjust the data so that the adjusted year means will be devoid of any effect of group. The first operation is that of setting the data up in a two way table such as that shown in Table 9. The data are divided into cow groups as the first step. A cow group consists of all individuals having identical experiences in that they have lactations in the same years. For instance Cow Group 1 contains 5 cows and the only cows in the herd who have records only in 1931 and 1932. In Cow Group 2 is one cow, the only one with records in 1931,1932, and 1933. Cow Group 3 consists of 2 cows having records in 1931,1932, 1933, and 1934. This continues to Cow Group 40 which contains 2 cows with lactations in 1944, 1945, and 1946. Altogether there are 40 cow groups containing 183 individual records. After the division is made according to cow groups, the records of each cow group are placed in their respective years. Thus we have subclasses divided on the basis of cow groups and years.

A simplification of the data thus set up may be provided. If a mean is taken for each cow group, and these means compared with the mean of the entire data, the differences show how much each group is above or below the population average. If the individuals within each group are corrected so that each group mean is equal to the over-all mean, then there are no differences between groups. This process is shown by the following equation:

$$(1) \quad X_{ij} - \bar{X}_j + \bar{X} = A_{ij}$$

where X_{ij} is the i th cow in the jth row or column, $\bar{X}_j$ is the mean of the jth row or column, $\bar{X}$ is the grand mean, and A_{ij} is the adjusted ith cow in the jth row or column. Then :

$$(2) \quad \frac{\sum (X_{ij} - \bar{X}_j + \bar{X})}{N_j} = \bar{X}$$

This does not affect the variability within subclasses; hence it is necessary to correct only the means of the subclasses:

TABLE IX. SAMPLE TWO WAY TABLE OF RECORDS

Cow Group	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	Mean	d
1	537	480															508.5	- 32.7
2	520	344	424														429.3	- 111.9
3	2 528	2 492	2 499	2 519													509.5	- 31.7
4	626	477	471	583	531												537.5	- 3.6
5	627		485	535	482												532.3	- 8.9
6	520	344	424		577												466.3	- 74.9
7	677	644	781		711	533											669.2	+ 128.0
8	331		695	633		717	802										635.6	+ 94.4
9	594	439	615	608		460	353										511.5	- 29.7
10		720	304														512.0	- 29.2
11		716	673	688													692.3	+ 151.1
12		546		589	495	618											562.0	+ 20.8
13			677	521	659												619.0	+ 77.8
14			542	594		421											519.0	- 22.2
15			724	654			509										629.0	+ 87.8
16			478	512	516	319	260										417.0	- 124.2
17			404	345		386	445		553								426.6	- 114.6
18					595	420	390	415									455.0	- 86.2
19						609	577										593.0	+ 51.8
20						573	642	610	602								606.8	+ 65.6
40														2 593	2 621	2 569	594.3	+ 53.1
No.	14	15	16	13	8	12	13	7	11	11	9	11	11	15	11	6	Grand Mean	541.2
Mean	545.4	507.6	543.4	561.5	570.8	506.8	527.5	554.4	572.0	561.5	512.1	564.4	514.4	526.8	529.3	624.8		

$$(3) \quad \bar{X}_{ij} - \bar{X}_j + \bar{X} = \bar{A}_{ij}$$

$\bar{X}_{ij}$ being the mean of the ith subclass in the jth row or column and $\bar{A}_{ij}$, the corrected mean of the ith subclass in the jth row or column. This is the process of adjustment to be used in the analysis.

The adjustment process is useful in this case where the subclass numbers are unequal. When subclass numbers are disproportional, the differences between border means are not true estimates of the parameter differences because the differences are determined not only by one classification but include some of the effects of the other classification. For example, the differences between year means contain some of the cow group differences; likewise group mean differences include some of the effects of year mean differences. The object then is to remove the effects of group mean differences from the year mean differences which may be done by adjusting according to the final equation.

The first step is to estimate the sum of squares due to year means. The year mean differences contain effects of group because of disproportionate subclass numbers; hence if all group means are corrected to the grand mean, the group effect will be removed from the year differences. This adjustment is shown in Table 10. All group means are adjusted to the grand mean so no difference remains between them. The sum of squares of years means with group effect removed is then 89,700. This figure is less than it should be because the removed group differences contained some year effects due to disproportionate subclass numbers; therefore an adjustment in the other direction must be made to remove these effects.

When year differences are removed by readjustment, the group

TABLE X. FIRST ADJUSTMENT FOR GROUP

Cow Group	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	Mean
1	5 570	5 513															541.2
2	632	456	536														"
3	2 560	2 524	2 531	2 551													"
4	630	481	475	587	535												"
5	636		494	544	491												"
6	595	419	500		652												"
7	549	516	653		583	405											"
8	237		601	539		623	708										"
9	624	469	645	638		490	383										"
10		749	333														"
11		565	522	537													"
12		525		568	474	597											"
13			599	443	581												"
14			564	616		443											"
15			636	566			421										"
16			602	636	640	443	384										"
17			519	460		501	560		668								"
18					681	506	476	501									"
19						557	525										"
20						507	576	544	536								"
40														2 540	2 568	2 516	"
Mean	562.1	519.4	546.2	556.5	579.7	502.4	523.7	533.8	572.2	556.5	517.0	558.8	524.1	520.8	539.7	577.6	Grand Mean - 541.2
d	+20.9	-21.8	+ 5.0	+15.3	+38.5	-38.8	-17.5	- 7.4	+31.0	+15.3	-24.2	+17.6	-17.1	-20.4	- 1.5	+36.4	

differences will no longer be zero, again due to unequal subclass numbers. The sum of squares is then calculated for group means yielding 8,200. This is a portion of the year sum of squares removed in the first adjustment for group and is added to the first sum 89,700 to give 97,900 as a sum of squares for between years.

Again readjustment must be made for group effects, and the weighted sum of squares calculated. As this procedure continues, each readjustment recovers a portion of the between year sum of squares lost in the first adjustment. The total of all the sums of squares is the sum of squares of between year means. This total consists of the contributions from both sets of border means. Continued readjustments result in smaller and smaller sums of squares, and they should cease when the adjustment results in only a fraction of the initial sum of squares.

The sum of squares for between years, in this case after 5 adjustments, is 89,700 + 8,200 + 1,200 + 1,600 + 700 or 101,400. Further adjustment in this direction is unnecessary because the part of the year differences unrecovered is small since the last sum of squares is only a fraction of the initial sum of squares.

The variation between years has been determined; the next step is to estimate the sum of squares for cow groups. This is accomplished by similar adjustments except that the first correction must be made to remove the effects of year this time. The same process follows and yields a sum of squares for between cow groups of 843,900 from 4 adjustments composed of sums of squares, 801,000 + 36,000 + 4,900 + 2,000.

The sums of squares for between years and between cow groups

having been determined, interaction can be calculated. The assumption is made that the sum of squares of the subclass means they are adjusted for border effects is an efficient estimate of the variance due to interaction. The combined results are shown in Table 11.

TABLE XI. ANALYSIS OF VARIANCE OF YEARLY DIFFERENCES IN PRODUCTION

Source of Variance	D/F	Sum of Squares	Mean Square
Between Subclasses	157	1,610,000	10,300
Between Year Means	15	101,400	6,760
Between Cow Group Means	39	843,900	21,638
F = 6,760/6,453 = 1.048			

The analysis of variance shows no evidence of significance in year mean differences. This may be due to the small number of individuals or the scattering and numerous subclasses, but no evidence of significance is demonstrated in the yearly environmental differences in this herd. One difficulty lies in the fact that there is more variance within year than between years.

Although the yearly differences in these data may not be significant, the rest of the method will be presented to compare with the two plans already presented. From the adjustments already made the adjusted year means, relatively free of group effect, may be determined. This consists of adding to the grand mean the correction factors or the differences between the succeeding group adjusted year means and the grand mean. Continually diminishing differences are obtained with each adjustment. For the purpose of this study the total of the differences between each group adjusted year mean and the grand mean may be used. The results of this process are shown in Table 12. From these the yearly differences can be computed by taking the differences between the total corrections of the years in question. To be on the same basis as the results of the previous trials, differences are taken between 1931 and each succeeding year.

As shown, this method has the advantages of utilizing all possible data and it has a test of significance. On the other hand, it requires considerable computation especially considering that at best the result is only an approximation.

TABLE XII. ENVIRONMENTAL DIFFERENCES DETERMINED BY ADJUST. METHOD

	Adjustments 1	2	3	* Total	Environmental Difference
1931	+ 20.9	+ 3.0	+2.2	+ 26.1	Base
1932	- 21.8	+ 0.9	+2.2	- 18.7	- 45
1933	+ 5.0	+ 3.1	+1.8	+ 9.9	- 16
1934	+ 15.3	+ 3.4	+1.6	+ 20.3	- 6
1935	+ 38.5	+ 6.9	+2.1	+ 47.5	+ 21
1936	- 38.8	- 5.7	-0.9	- 45.4	- 72
1937	-17.5	- 4.9	-1.6	- 24.0	- 50
1938	- 7.4	- 4.3	-2.0	- 13.7	- 40
1939	+ 31.0	+ 1.6	-0.6	+ 32.0	+ 6
1940	+ 15.3	+ 2.5	+0.5	+ 18.3	- 8
1941	- 24.2	+ 1.9	+0.8	- 21.5	- 48
1942	+ 17.6	+ 1.3	-0.7	+ 18.2	- 8
1943	- 17.1	- 3.2	-1.7	- 22.0	- 48
1944	- 20.4	- 3.3	-2.0	- 25.7	- 52
1945	- 1.5	- 4.9	-2.7	- 9.1	- 35
1946	+ 36.4	+ 3.1	-0.9	+ 38.6	+ 13

*Sum of group adjusted year mean - grand mean differences for three group adjustments.

COMPARISON OF PLANS

Three methods have been presented for determining the yearly environmental differences within a herd. It might now be desirable to compare these methods as to their possible merit.

Each of the methods was used to calculate the yearly environmental differences in the M.S.C. Holstein herd over a period of 16 years. To be placed on an equivalent basis, they were computed in relation to the base or first year, 1931. Table 13 contains the results of the three plans including 5 variations of the second plan. It must be remembered that the number of records was small which limits the validity of any of the calculations, and the analysis of variance gave no indication of significant year differences in average production.

The adjusted means method supposedly gives the most likely correct values for yearly environmental differences although the deviations are a little smaller than they might be if the adjustments had been carried further. Since this method should be reliable, it is used as a standard for comparison with the less elaborate systems.

I-A, the paired year comparison, is apparently completely out of proportion in this study. There is an extreme accumulation of negative values throwing the comparison very much out of line. I-B, likewise, is not well in agreement with the other results. Why this system should produce such an increasing negative value still is not apparent.

II-A, II-B, and II-C are the variations of the direct comparison plan utilizing all the comparisons with every year. As was stated earlier, there is some question as to whether the weighted average was really justified in this case where a direct comparison is more reliable than an

TABLE XIII. ENVIRONMENTAL DIFFERENCES DETERMINED BY VARIOUS METHODS

*	I-A	I-B	II-A	II-B	II-C	II-D	II-E	III
1931				Base Year				
1932	- 79	- 50	- 79	- 79	- 50	- 79	- 50	- 45
1933	- 86	- 37	- 41	- 41	- 3	- 58	- 6	- 16
1934	- 91	- 46	- 22	- 14	+ 10	- 36	+ 6	- 6
1935	- 103	- 66	- 23	- 21	+ 11	- 54	+ 5	+ 21
1936	- 212	- 167	- 96	- 83	- 60	- 55	- 35	- 72
1937	- 189	- 147	- 98	- 104	- 71	- 55	- 25	- 50
1938	- 196	- 132	- 98	- 120	- 54	- 47	- 1	- 40
1939	- 179	- 65	- 42	+ 19	- 1	+ 25	+ 53	+ 6
1940	- 191	- 72	- 71	- 59	- 11	+ 20	+ 41	- 8
1941	- 235	- 123	-116	- 97	- 58	- 28	- 8	- 48
1942	- 230	- 107	- 80	- 71	- 20	+ 20	+ 39	- 8
1943	- 283	- 160	-124	- 114	- 64	- 34	- 14	- 48
1944	- 288	- 166	-133	- 128	- 75	- 39	- 20	- 52
1945	- 278	- 154	-119	- 96	- 66	- 32	- 11	- 35
1946	- 293	- 150	- 48	- 21	+ 28	+ 21	+ 69	+ 13

* I-A Paired Year Comparisons

I-B Paired Year Comparisons (Regression Corrected)

II-A Direct Comparison (Weighted)

II -B Direct Comparison (Unweighted)

II-C Direct Comparisons (Regression Corrected)

II-D Direct Comparisons (Non-Duplicating)

II-E Direct Comparison (Regression Corrected) (Non-Duplicating)

III Adjusted Means

indirect comparison, but if it is assumed that the indirect comparisons are reliable, the weighted average is desirable. II-C is theoretically the best estimate of the three as it includes the weighted averages and is also corrected to counteract the effects of the expected regression from a single record toward the herd average in succeeding records. It will be noted that it is in general agreement with the "adjusted mean" values. The trend in direction in every year except 1937 agrees, or each year is + or - in respect to the preceding year in both results.

II-D and II-E are the products of the same type of calculation as the preceding three methods, but each individual is used in only one comparison rather than being used repeatedly as in the preceding types. II-D is the straight weighted average calculation while II-E is corrected for regression. In both of these the general picture is similar to the "adjusted means" result.

One point demonstrates itself fairly well in that the regression corrected figures are drawn closer to zero deviation than the uncorrected. Where negative differences predominated in the uncorrected figures, the regression correction pulls them down until positive differences appear also. The regression corrected figures of all methods correspond more closely to the "adjusted mean" results than do the uncorrected.

If the "adjusted mean" differences are assumed to be the most reliable, the most closely agreeing would be method II-C or II-E. The numbers of records are so small, however, that no definite statement should be made concerning the best system especially in view of the fact that no significant difference between years was evident.

It seemed that it might be interesting to apply these differences as corrections to the actual herd averages each year to see what picture they might produce. Each difference was applied to the herd average of each respective year as a correction or with its sign reversed. If the yearly environmental difference was positive, the year was better than the base year; hence, that difference must be deducted from the herd average to make all yearly averages equivalent. The results of applying I-B and II-A are shown graphically in comparison with III and the actual herd averages in Figure 3; II-C and II-E are compared in Figure 4.

With the assumption that III is correct, its curve would indicate the genetic trend of the herd. By means of the computation of Plan III nearly all environmental differences are removed from the herd average leaving only the heredity differences. This then gives the picture of the success of the breeding and selection operations carried on by the breeder in respect to butterfat production.

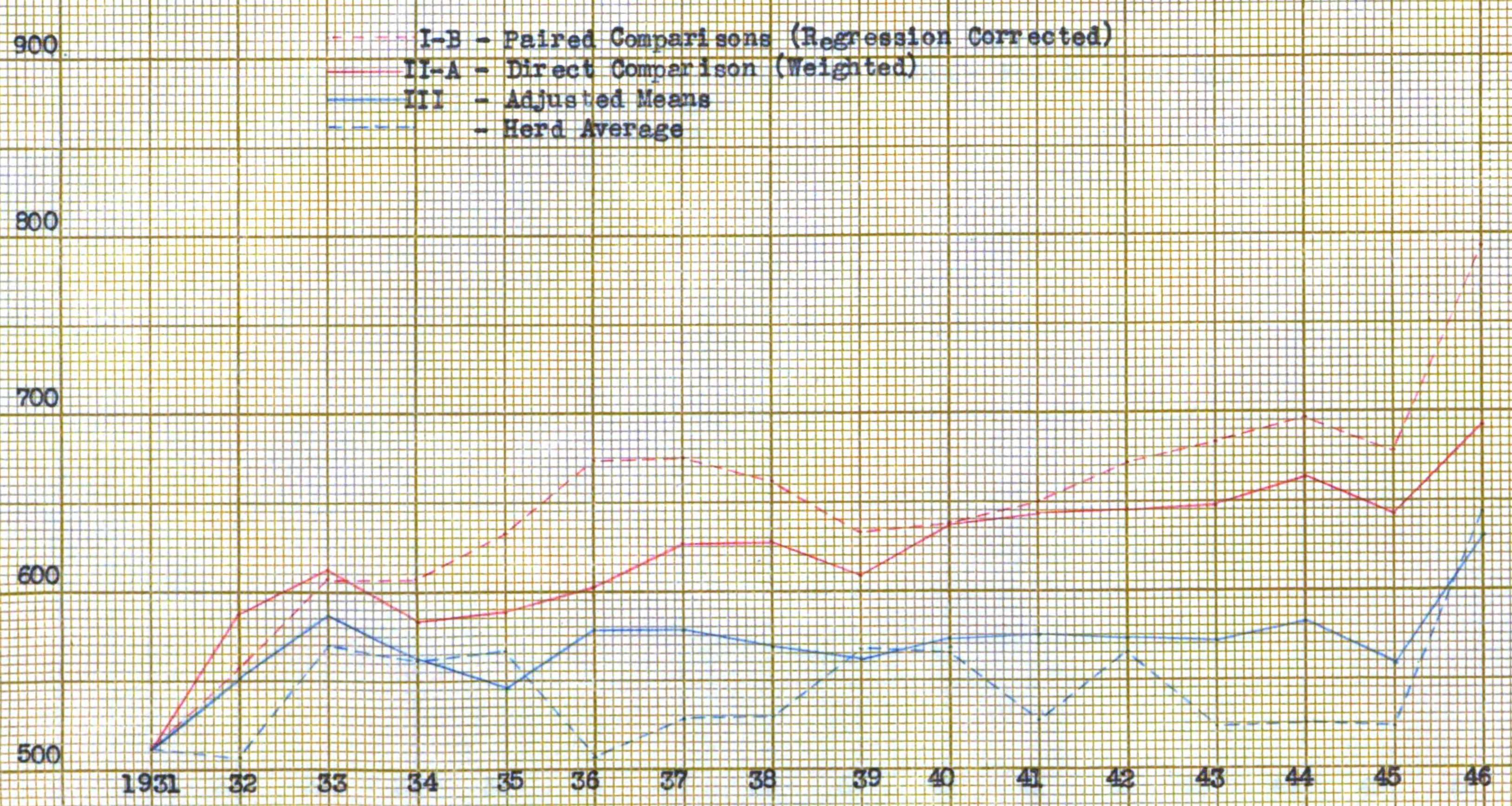

FIG. 3. COMPARISON OF CORRECTED HERD AVERAGES

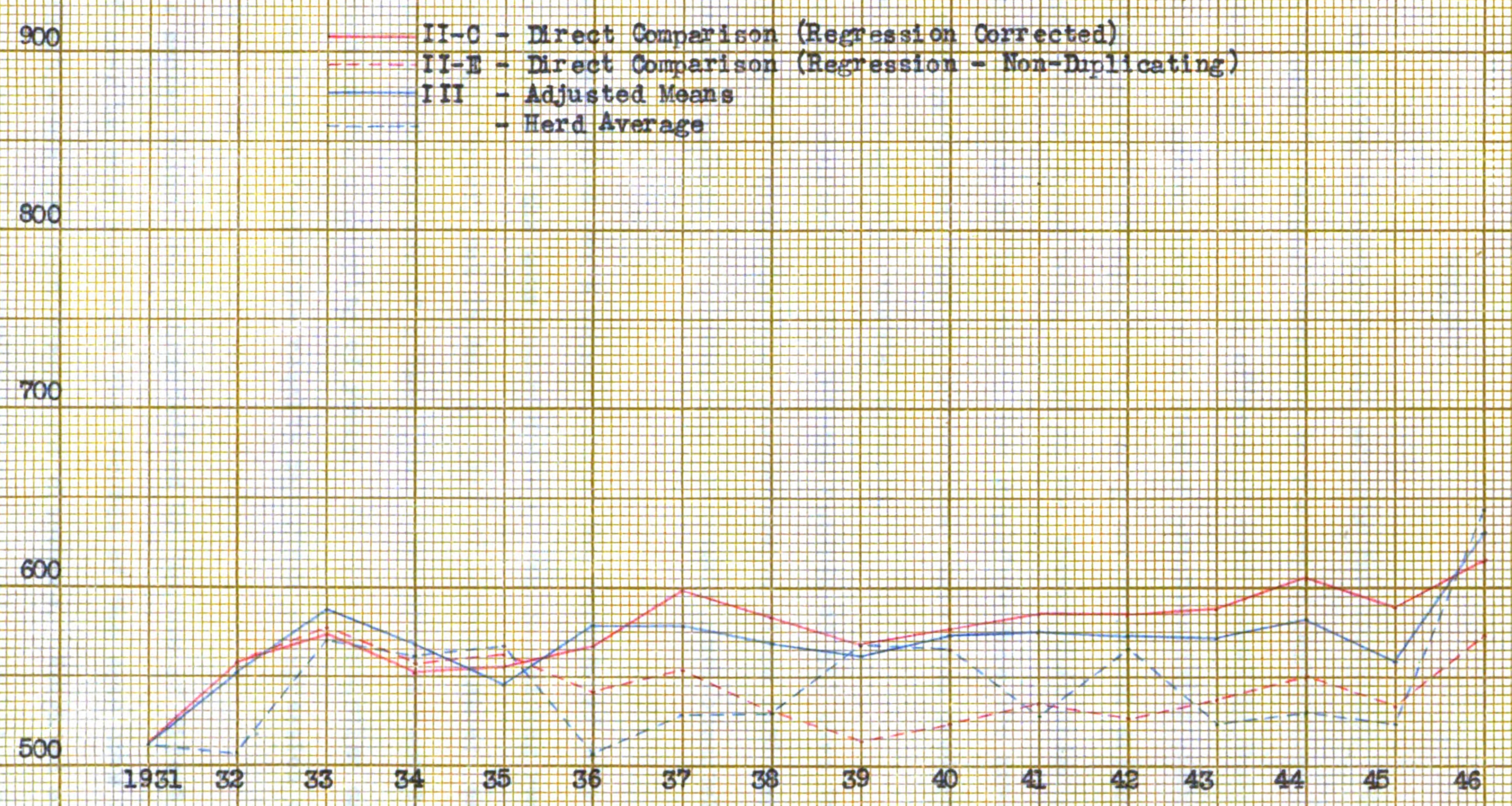

FIG. 4. COMPARISON OF CORRECTED HERD AVERAGES

SUMMARY

The mistaking of environmental effects for hereditary effects is one of the breeder's largest barriers to the rapid improvement of his herd. The removal of the confusing effects of environment is highly desirable and can be accomplished with greater practicability statistically.

Various methods of equalizing yearly environmental differences were presented, all of these being based on the theory that under constant environmental conditions, an individual cow should produce the same amount of butterfat each lactation. Any variation from this even production should be due to varying environmental effects.

The mature equivalent, 3X, 305-day records of all cows in the Holstein herd at M.S.C. with more than one lactation since 1930 used to compare the various methods of determining the yearly environmental differences in the herd.

The first method was a comparison of the records of cows in one year with the records of the same individuals in the succeeding year. This method is simple and brief, but it has no test of significance; any errors may be cumulative with the use of a base year; and not all records can be utilized where there is a skipping of years. In this particular case selection in the herd affected the results for which a regression correction was made.

The second method was characterized by the use of direct comparison of records of cows in each year with the records of the same cows in a base year where possible and indirect comparison elsewhere. Weighted and straight averages for computing difference values and regression corrections

were utilized and compared. Another variation was presented in which duplications of comparisons were avoided and the comparisons of records of the same cow were used only once. This general method is slightly more complicated than the first method, but it has the added advantage of utilizing all available records of cows with two or more lactations. There is still no test of significance.

The final method was a lengthy process of adjustment of border means. The computation is the limiting factor in this method. It does have the advantage of utilizing all of the records, and there is a test of significance to indicate if there is a yearly difference which needs correcting.

Since there was no evidence of significant difference between years, and since the number of records was small, few conclusions may be drawn from the analysis of the methods. They are suggested as possibilities. Where selection is employed in the herd as is the usual occurrence, some correction must be made. A more precise correction than the regression factor used in this study must be determined.

There is considerable question as to the validity of mature equivalent conversion factors which may have some bearing on this work. These must be correct or the entire basis of the determination is destroyed. There is some question as to how useful they are, and there is no doubt but that they may be very incorrect for certain individuals. When there is such a definite doubt about M.E. conversion factors, more errors may be introduced than are removed. Probably the only solution to this is to use only the actual mature records of the cows. Obviously this would require a tremendous number of cows in the herd, more than the average herd contains since the make-up of most herds is young cows.

It would seem that the type of environmental corrections discussed in this paper are, in general fairly impractical at present. If only records of mature cows in sufficient numbers were available, and if it could be determined what their likely production might be to eliminate the regression factor or if they were unselected animals, there might be encouragement for the use of one of these methods. At present the value obtained from the results may be overbalanced by the error and the computation involved.

LITERATURE CITED

(1) Johannson, I.
Undersokningar över avkastnings - resultaten inom thorsätrabesättningen under stallfodring och betsesgång. Noddelande Från Kungl. Landtbruksakademiens Lantbruksavdelning Nr. 13: 1-53. 1938
(Cited by R. H. Nelson. See Reference (3)).

(2) Lush, Jay L.
Animal Breeding Plans, Third Edition.
Iowa State College Press. 1945.

(3) Nelson, R. H.
The Effects of Inbreeding on a Herd of Holstein-Friesian Cattle.
Thesis for Ph.D. degree, Iowa State College. 1943.

(4) Patterson, R. E.
The Use of Adjusting Factors in the Analysis of Data with Disproportionate Subclass Numbers.
Jour. Am. Stat. Assoc., 41: 334-346. 1946.

(5) Yates, F.
The Analysis of Multiple Classifications with Unequal Numbers in the Different Classes.
Jour. Am. Stat. Assoc., 29: 51-69. 1934.