PREDICTION OF LACTATION YIELD OF A COW FROM HER PARTIAL YIELD.

By

T. V. SUNDARARAMAN

Institute of Agricultural Research Statistics (I. C. A. R.)

Library Avenue, New Delhi.

1 9 6 0.

PREDICTION OF LACTATION YIELD OF A COW FROM HER PARTIAL YIELD.

By

T. V. SUNDARARAMAN

Dissertation

submitted in partial fulfilment

of the requirements for the award of Diploma

in Agricultural and Animal Husbandry S t a t i s t i c s

of the Institute of Agricultural Research Statistics (I.C.A.R.)

New Delhi.

ACKNOWLEDGEMENTS.

I have great pleasure in expressing my deep sense of gratitude to Dr. V.N. Amble, Assistant Statistical Advisor (Animal Husbandry), Institute of Agricultural Research Statistics (I.C.A.R.), for his valuable guidance, keen interest and constant encouragement throughout the course of investigation and of preparation of the 'thesis'.

I am highly thankful to Dr. V.G. Panse and Dr. G.R. Seth, Statistical Advisor, I.C.A.R., for providing me with adequate facilities for my work.

My thanks are also due to Dr. P.N. Saxena, Chief Statistician, I.A.R.I., for making available the results of the unpublished study on the Sahiwal herd at the I.A.R.I.

T.V. Sundararaman.

(T.V. SUNDARARAMAN.)

CONTENTS.

1. INTRODUCTION.

The prediction of the lactation yield at an early stage is very useful for a dairy man in that it enables him to know what the expected production of his stock would be. He wishes to know, when an animal has been in milk for only a part of the lactation, what its total lactation yield would be so that if it is an uneconomical animal, he may dispose it of without incurring further, unnecessary, expenditure on it. In herds maintained for raising breeding stock it is equally profitable to have a precise method of predicting the lactation yield, as for example, in deciding on the male calves to be retained for possible breeding use on the basis of the dams' lactation performance. Such selection, if it could be safely made early without awaiting the completion of the lactation by the dams, could result in the considerable saving in the expenditure on raising the unwanted male calves. If partial records were found to be sufficiently reliable indicators of total lactation yields, they would also be of some help in cutting down the period required for evaluating the breeding worth of a Sire. They could similarly be helpful in shortening the period for carrying out experiments of some types on dairy cattle. Thus a quick, easy and accurate method of estimating the lactation yield would be useful to commercial dairy men and animal breeders alike.

The object of the work reported in the present 'thesis' was to examine the suitability of the use of partial lactation records as predictors of lactation yields on the basis of data on Indian herds of dairy cattle.

2. REVIEW OF LITERATURE.

Many workers have studied the problem of predicting lactation yield from partial yield.

Caines (1927) found, on analysing data on Holstein cows, that since a seven-day test just after calving does not depend upon and does not measure persistency, the short-time test should be conducted during the fourth month of lactation to best represent the practically useful lactation capacity of the cow.

Tuff (1928) found that total lactation yield of a cow could be obtained by multiplying her maximum daily yield by 199.

Sanders (1930), in an attempt to analyse lactation yield into its two components, namely, persistency and maximum yield, found that the ratio of total lactation yield to the maximum yield varied between 24.7 and 29.87 according to the month of calving.

Kartha (1934), in a study of data on specially bred Indian cows, buffaloes, cross-bred cows and ordinary Indian cows, correlated each month's yield, taken separately, with total yield and observed that while the fourth month's yield gave the highest correlation coefficient of +0.9333, the yields in the 3rd, 5th and 6th months gave fairly high coefficients. He also observed that the peak of lactation was generally reached between the 3rd and 6th weeks after calving and that the average daily yield for these four weeks multiplied by a constant factor gave the yield for 300 days. He arrived at different constant factors for the different breeds.

Gooch (1935), in a study of the monthly production records of pure-bred Jerseys, correlated, among other things, milk yield for the first month with the total lactation yield and concluded that the amount of milk that a cow gives at the beginning of her lactation predicts, quite effectively, her ultimate worth as measured by her total lactation yield.

McGilliard and Smith (1941) made a comparison of different methods of calculating yearly milk and butterfat records such as tests taken at weekly, monthly, or bi-monthly intervals. They made their study on 100 yearly Holstein and Jersey records and concluded that only small differences existed between accuracy of different methods.

Cannon, Frye and Sims (1942) observed that the prediction of total yield based on regression is most accurately made from a test taken during the fifth month of lactation because of its low standard error.

Patow (1949) in his report of a paper read by Lenz quoted Lenz as saying that the lactation yield could not be accurately estimated from the yield of the period of 5 to 34 (30 days) after calving.

Mahadevan (1951) studied the interrelation-ships of persistency, initial milk yield for the first ten weeks and 180 days' milk yield from a large body of data on Ayrshire cattle in Scotland and observed that the three characteristics were positively correlated and that the correlation between first 10 weeks' yield and the total yield for 180 days was highly significant.

Bussert (1957) studied the relationships between partial yields, based on 100- and 200- day yields, and total lactation yield, taking into consideration the shape of the lactation curve. He found that while the 100-day yield was not an accurate guide to later production, the 200-day yield was suitable for estimating milk producing ability in heifers and did not need to be considered in relation to the shape of the lactation curve.

Rendel et.al(1957) observed that 70-days yield had a correlation coefficient of +.80 with the total lactation yield.

Bognoni and Brambilla (1957) found that the correlation coefficient between milk yield during the first 70 days of lactation and the total lactation yield, in addition to being high, was, in general, greater than the correlation coefficient between milk yield during the first 280 days and the total lactation yield.

Amble, Rajagopalan and Gill (1958) made a study on 'sampling of daily milk records for estimating lactation yield of cows' from data on six Government cattle farms and found that systematic sampling, in addition to being convenient in practice, was more efficient than simple random sampling, whatever be the interval of recording, and that the recording of milk yield at weekly intervals would provide an estimate of a single lactation yield with a precision corresponding to a standard error of 1.8 to 2 per cent.

Saxena and Kumar analysed data belonging to the first two lactations of the Sahiwal herd at the Indian Agricultural Research Institute and found that correlation coefficient between 8 weeks' yield and total yield was .83 for the first lactation and .72 for the second.

3. EXTENT AND NATURE OF DATA.

In order that the inferences drawn may be broad-based, data pertaining to seven different populations of Indian cattle were taken for study. The data included the daily yield records of Red Sindhi cows at the Livestock Research Station at Hosur, Red Sindhi cows at the then Indian Dairy Research Institute at Bangalore, Tharparkar cows at the Government Cattle Farm, Patna, Hariana cows at the Indian Veterinary Research Institute, Izatnagar, Hariana cows maintained under a scheme of the Indian Council of Agricultural Research at the Government Livestock Farm at Hissar and of the cows in the selected households under the survey for the estimation of cost of milk production in 1953-55 conducted by the I.C.A.R.

In the Tharparkar herd at Patna, two types of data were collected; in one, calves were weaned at birth and in the other calves were not so weaned. These two data were treated separately. Among the other herds, calves were weaned at birth in the Red Sindhi herd at Bangalore and in the Hariana herd at Izatnagar while in the remaining herds, weaning was not practised.

Since the lactation yield varies from lactation to lactation it would be necessary to base comparisons between groups of cows, such as are involved in Sire evaluation, on the performance in the same lactation. For such purposes the first lactation being the earliest is to be preferred. For this reason it was decided to confine the study to the first lactation.

The nature and extent of data which were available are indicated in table I.

<u>Table I.</u>

Details regarding the herds examined under the study.

No.	Name of the herd or group.	Weaning practice.	No. of cows.	Average lact. length (weeks).	Average lact. yield (lb.)
1.	Red Sindhi, Hosur.	Not weaned.	34	44	2148
2.	Red Sindhi, Bangalore.	Weaned.	40	41	3305
3.	Tharparkar, Patna.	Weaned.	65	44	3233
4.	Tharparkar, Patna.	Not weaned.	38	47	2650
5.	Hariana, I.V.R.I.	Weaned.	33	45	2610
6.	Hariana, Hissar.	Not weaned.	35	43	1201
7.	Cost of milk production survey, Delhi.	Not weaned.	31	33	1383

In calculating the lactation yield and the lactation length the period of colostrum was excluded. This period varied from two to nine days in the case of Hariana cows in the Hissar Farm. For the other herds the first seven days were omitted. In the case of herds where weaning was not practised, on one day in a week the milk was completely stripped. A study of the stripped yields had been made in an earlier study at the Institute and it was found that they followed the same trend as the corresponding unstripped yields but differed only in magnitude by an amount which was more or less constant all through the lactation. For obvious reasons the study was confined to unstripped yields which were six times as numerous in such herds. The stripped yield on the seventh day was substituted by the average of the unstripped yields of three days preceding and three days following the day in order to obtain a smooth lactation curve.

The data on Delhi cows were studied in order to see to what extent the results in the case of Government Livestock Farms compare with the results based on records of animals of private owners brought up under the conditions of feeding and management prevailing in rural areas.

4. PROCEDURE.

4.1. Correlation studies.

Let z denote the total lactation yield of a cow given by her yield for the first 300 days or less of lactation and let x denote her partial yield for an initial period of w weeks (w = 1, 2, 3 ....). $y = z - x$ is then the cow's subsequent or later yield. The problem is to set up a prediction equation predicting the total lactation yield of a cow from her partial yield. Let the expected lactation yield on the basis of prediction equation be denoted by Z. Let n be the number of cows in a herd.

To start with, for each of the herds, scatter diagrams (Figures 1.) were drawn between partial yields x and later yields y to roughly find out if the relationship between x and y was linear. There was a good indication of linearity in trend. The number of cows in each set of data was not large enough to permit suitable grouping of the data by confirming the trend by carrying out a test of linearity. correlation coefficients r_{xz} between x and z and the like coefficients r_{xy} between x and y were then worked out for x yields taken upto the first eight, twelve and sixteen weeks and also upto the peak week. It may be mentioned that r_{xz} corresponding to x yield for less than eight weeks was not tried because of the apparent none too high values for r_{xz} even for 8 weeks' yield.

It is desirable that the prediction equation of linear regression should explain atleast eighty per cent. of the total variation in Z. Since the percentage variation in Z explained by its linear regression on x is just the square of the correlation r_{xz}, eighty per cent of the regression variation would correspond to a value of the order of .90 for the correlation. The study was hence extended to partial yields of longer duration until the shortest period (in weeks) for each herd was secured such that the correlation between the yield upto that period and the total lactation yield was +.90.

4. 2. Prediction equation.

1. Linear regression of Z on x is represented by the st. line equation $Z = \alpha + \beta x$, α and β being the coefficients. The theory of linear regression involved in the estimation of α and β is briefly described below:

First, the observed value (z) of Z for any given x is regarded as a random sample from a hypothetical frequency distribution which could be generated by repeating the measurement process over and over again at this same x. The true mean value of this distribution or the expected value Z is the quantity that is linearly related to x. The actual individual observations, obviously, differ from it. Thus each observation is of the form $Z = \eta + e$ where η is the true value $\alpha + \beta x$ and $e = Z - \eta$ is the random residue (error). 'Least squares' method of estimation gives the minimum variance unbiased estimates of α and β. Here it is assumed that the average value of the errors is zero and they have a common variance (σ_e^2). If a and b respectively denote the estimates of the parameters α and β, the principle of 'least squares' consists in choosing 'a' and 'b' such that $\sum_1^n (z-Z)^2$ is minimum, n being the number of sample observations. The values of a and b thus obtained can be easily shown to be

$$b = \frac{\sum_{1}^{n} (x-\bar{x})\,(z-\bar{z})}{\sum_{1}^{n} (x-\bar{x})^2} \qquad \text{———————— (1)}$$

and $$a = \bar{z} - b\bar{x} \qquad \text{———————— (2)}$$

where $$\bar{x} = \frac{\sum_{1}^{n} x}{n} \quad \text{and} \quad \bar{z} = \frac{\sum_{1}^{n} z}{n}$$

The linear equation thus obtained is used as the prediction equation for predicting the lactation yield Z from the partial yield x. From the formulae (1) and (2), 'a' and 'b' and hence the prediction equation were calculated for each of the herds.

2. <u>Standard error of 'b'</u>

The error sum of squares $\sum_{1}^{n}(z-Z)^2$ is

$$\left[\sum_{1}^{n}(z-\bar{z})^2 - \frac{\sum_{1}^{n} (x-\bar{x})\,(z-\bar{z})}{\sum_{1}^{n}(x-\bar{x})^2} \right].$$

If s_e^2 denotes the residual mean square, then

$$s_e^2 = \left[\sum_{1}^{n}(z-\bar{z})^2 - \frac{\sum_{1}^{n} (x-\bar{x})\,(z-\bar{z})}{\sum_{1}^{n} (x-\bar{x})^2} \right] \Big/ (n-2), \quad n-2$$

being the degrees of freedom for the error sum of squares.

Since $b = \frac{\sum_{1}^{n} (x-\bar{x})(z-\bar{z})}{\sum_{1}^{n} (x-\bar{x})^2}$, the variance of

of b would be $\frac{\sigma_e^2}{\sum_{1}^{n} (x-\bar{x})^2}$, σ_e^2 being

the dispersion of z around the regression line. is

estimated by the error mean square s_e^2 so that the standard error

of b obtained from the sample is

$$\left[\frac{s_e^2}{\sum_{1}^{n} (x-\bar{x})^2} \right]^{\frac{1}{2}} \quad \text{———} (3).$$

Using this formula (3), standard error of b was

worked out for each of the prediction equations.

4. 3. Accuracy in prediction of lactation yield.

1. As pointed out earlier, the linear equation $Z = a + bx$ gives the lactation yield Z of a cow from her partial yield x. It is not safe to use this prediction equation if the accuracy with which Z is predicted (from x) is low, the accuracy being measured by the standard error of the predictor.

The predicted value Z is subject to two sources of error. The first one is made up of the sampling errors in a and b, being respectively the estimates of α and β. a can be written as $\alpha + e_1$ and b as $\beta + e_2$ where $E(e_1) = 0 = E(e_2)$, $V(e_1) = V(a)$, $V(e_2) = V(b)$ and cov. $(e_1, e_2) = 0$.

The second error arises out of the fact that the expected value of $a + bx$, namely, $\alpha + \beta x$ gives only the average or the expected value of Z and not any individual value, an individual value being $\alpha + \beta x + e$.

The first error may be called the error of the prediction equation. The second error gives the residual variation in Z from its expected value.

If the prediction equation is used to estimate the true average value of Z, for a given x_0 of x, the variance of the predicted value is due to only the first source; the residual variance is absent here since the expected value of $a + bx \ (= \bar{z} + b(x - \bar{x}))$ itself gives the average or expected value of $z \{ E(e) = 0 \}$. The variance is hence

$$\left[\sigma_e^2 \left\{ \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_1^n (x - \bar{x})^2} \right\} \right]$$ being the sum of the variances

due to $\bar{z}$ and b.

If, however, the prediction equation is used to predict any individual value z, given the value x_0 of x, the deviation in z from its expectation which is what the prediction equation gives also creeps in and the variance of the predictor is the sum of the two variances and, since the residual variation is σ_e^2, the variance of the predictor is

$$\left[\sigma_e^2 \left\{ 1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_1^n (x - \bar{x})^2} \right\} \right]$$

Hood (1950) has generalised the above result for the average of m values. In this case, the residual variation is $\frac{\sigma_e^2}{m}$ $\left[\text{since } V(e) = \sigma_e^2, \ V(\bar{e}_m) = V \left\{ \frac{\sum_1^m e}{m} \right\} = m \frac{\sigma_e^2}{m^2} \right]$ so that the variance of the predictor is

$$\left[\sigma_e^2 \left\{ \frac{1}{m} + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum (x - \bar{x})^2} \right\} \right] ; \sigma_e^2 \text{ is estimated by the}$$

residual mean square s_0^2.

Thus the standard error of predictor giving the average lactation yield Z_0 of a group of m cows with an average initial yield x_0 is

$$\left[s_0^2 \left\{ \frac{1}{m} + \frac{1}{m} + \frac{(x_0 - \bar{x})^2}{\sum (x - \bar{x})^2} \right\} \right]^{\frac{1}{2}} \qquad \text{———— (4)}$$

The standard error thus depends, on the one hand, on the number of observations and the variability in the early yield values on which the prediction equation is based and, on the other, on the average initial yield x_0 and the size of the group m. The standard error increases as x_0 moves away from $\bar{x}$ and it goes on diminishing with an increase in the value of m. In order to obtain a fair idea of the accuracy in predicting the total lactation yield Z from its regression on partial yield x, a set of values of x_0 on either side of $\bar{x}$ and covering a wide range of the observed initial yields was selected with a suitable number of m ranging from unity onwards.

2. The 95 per cent. fiducial bands for the prediction line $Z_0 = \bar{z} + b(x_0 - \bar{x})$, predicting the average lactation yield of a group of m cows, are $Z_0 \pm t$ S.E. (Z_0) where t is the "Student's t" with n-2 degrees of freedom and for probability .05, and S.E.(Z_0) is the value

$$\left[S_e^2 \left\{ \frac{1}{m} + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\sum_1^n (x - \bar{x})^2} \right\} \right]^{\frac{1}{2}}.$$

Fiducial bands were obtained for the extreme values of m = 1 and m = ∞ ; bands corresponding to other values of m will lie within the range covered by these two sets. The fiducial bands are represented in Figures 2.

4.4. Curvilinear regression.

Even though the scatter diagrams indicated the relationship between the yield in the initial period and the later yield to be roughly linear, the possibility of improving the fit further by employing curvilinear regression was explored to find out whether or not curvilinear regression in x_1 of a certain initial period could explain as much variation as was explained by linear regression in x_2 of a greater initial period, x_2 being obviously greater than x_1. In such a case since prediction of lactation yield could be done earlier with its curvilinear regression on partial yield for a shorter duration, it would be advisable to prefer it even though fitting of curvilinear regression equations may be slightly more complicated than fitting of linear regression equations.

With a view to examining this, quadratic and cubic curves were fitted to data of 12- and 16-week yields in the case of four of the herds. Though the 'least squares' procedure of fitting quadratic and cubic curves is a known one, yet it is, briefly, discussed here for the sake of continuity.

Let the cubic equation (for regression of Z on x) in x's be denoted by

$$Z = \alpha + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 .$$

The estimates of β's, the regression co-efficients, are obtained by minimising the residual sum of squares $\sum_1^n (z-Z)^2$.

Minimisation leads to, after modification, the following equations

$$b_1 \sum(x-\bar{x})^2 + b_2 \sum(x-\bar{x})(x^2-\overline{x^2}) + b_3 \sum(x-\bar{x})(x^3-\overline{x^3}) = \sum(x-\bar{x})(z-\bar{z})$$

$$b_1 \sum(x-\bar{x})(x^2-\overline{x^2}) + b_2 \sum(x^2-\overline{x^2})^2 + b_3 \sum(x^2-\overline{x^2})(x^3-\overline{x^3}) = \sum(x^2-\overline{x^2})(z-\bar{z})$$

$$b_1 \sum(x-\bar{x})(x^3-\overline{x^3}) + b_2 \sum(x^2-\overline{x^2})(x^3-\overline{x^3}) + b_3 \sum(x^3-\overline{x^3})^2 = \sum(x^3-\overline{x^3})(z-\bar{z}),$$

where

$$\bar{z} = \frac{\sum z}{n},\ \bar{x} = \frac{\sum x}{n},\ \overline{x^2} = \frac{\sum x^2}{n} \text{ and } \overline{x^3} = \frac{\sum x^3}{n},$$

all summations being from 1 to n; b_1, b_2 and b_3 are respectively the unbiased estimates of β_1, β_2 and β_3.

If the three solutions of the successive equations are denoted by

$$\begin{matrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{matrix} \quad \text{with } c_{ij} = c_{ji}\ (i, j = 1, 2, 3.),$$

then

$$b_1 = c_{11}\Sigma(x-\bar{x})(z-\bar{z}) + c_{12}\Sigma(x^2-\bar{x}^2)(z-\bar{z}) + c_{13}\Sigma(x^3-\bar{x}^3)(z-\bar{z}).$$

$$b_2 = c_{21}\Sigma(x-\bar{x})(z-\bar{z}) + c_{22}\Sigma(x^2-\bar{x}^2)(z-\bar{z}) + c_{23}\Sigma(x^3-\bar{x}^3)(z-\bar{z}).$$

$$b_3 = c_{31}\Sigma(x-\bar{x})(z-\bar{z}) + c_{32}\Sigma(x^2-\bar{x}^2)(z-\bar{z}) + c_{33}\Sigma(x^3-\bar{x}^3)(z-\bar{z}).$$

a, the estimate of α, is then obtained from

$a = \bar{z} - b_1\bar{x} - b_2\bar{x}^2 - b_3\bar{x}^3$ so that the prediction equation

becomes

$$Z = \bar{z} + b\,(x-\bar{x}) + b\,(x^2-\bar{x}^2) + b_3\,(x^3-\bar{x}^3).$$

The residual variation $\Sigma(z-Z)^2$ becomes, after simplification,

$$\left[\left\{\Sigma(z-\bar{z})^2\right\} - \left\{b_1\Sigma(x-\bar{x})(z-\bar{z}) + b_2\Sigma(x^2-\bar{x}^2)(z-\bar{z}) + b_3\Sigma(x^3-\bar{x}^3)(z-\bar{z})\right\}\right].$$

The term in the first bracket is the total variation in Z while the terms in the second constitute the variation in Z due to its regression on x. Hence, out of the total variation $\Sigma(z-\bar{z})^2$, the fraction which is explained by the third - degree regression curve $Z = a + b_1x + b_2x^2 + b_3x^3$ is

$$\left|\frac{\left\{b_1\Sigma(x-\bar{x})(z-\bar{z}) + b_2\Sigma(x^2-\bar{x}^2)(z-\bar{z}) + b_3\Sigma(x^3-\bar{x}^3)(z-\bar{z})\right\}}{\left\{\Sigma(z-\bar{z})^2\right\}}\right| \qquad \text{———}(6)$$

Similarly, the fraction explained by a quadratic can be worked out.

4. 5. Prediction of Average Lactation Yield of a group of cows from a sample.

For scientific animal breeding it is quite important to know the number of daughters needed to establish as significant the breeding worth of a bull whose index may turn out to be greater than a given herd value by a specific percentage of the value. More generally, it may be desired to establish the significance of the superiority of a group of N' cows on the basis of records of a random sample m' of the cows. It is of interest to investigate the increase in the number of cows required to be sampled if the test of significance is to be based on their partial records.

Let I denote the estimated simple daughter average of a bull based on the records of m' daughters. Let K be the percentage superiority of the bull over the given herd average A at 5 per cent. level of significance.

Since m' cows are a random sample from among N' cows,

$V(I) = \left\{ \frac{1}{m'} - \frac{1}{N'} \right\} S_b'^2$ where $S_b'^2$ is the

mean square between lactation yields of the N' cows. In practice, however, $S_b'^2$ is not known and, as such, is to be substituted by its estimate $s_b'^2$, being the mean square between lactation yields of the m' cows. The ratio

$$t = \frac{(I - A)}{\left[\left\{\frac{1}{m} - \frac{1}{M}\right\} s_b'^2\right]^{\frac{1}{2}}}$$ is distributed

as a "Student's t" with m' - 1 degrees of freedom; the test is a one - tailed one.

Since,

$$I - A = \frac{A \times K}{100}, \text{ ———————— } (6)$$

$$\frac{A \times K}{100} = t \left[\left\{\frac{1}{m'} - \frac{1}{M'}\right\} s_b'^2\right]^{\frac{1}{2}}$$ and hence

$$m' = \frac{1}{\left\{\frac{A^2 K^2}{(100)^2 s_b'^2} + \frac{1}{M'}\right\}} \text{ ———————— } (7)$$

In the most important case of M' = ∞, that is when the sample of m' is to be drawn from a population which is conceptually infinite in magnitude, the equation (7) becomes

$$\frac{1}{\left\{\frac{A^2 K^2}{(100)^2 s_b'^2 t'}\right\}} = \frac{t^2 s_b'^2 (100)^2}{K^2 A^2}.$$

Different values for K and M' were used in equation (7) and the corresponding values for m' were obtained, iteration being necessary since the degrees of freedom of t depends on m'.

If the lactation yields Z_0 of these m' cows are predicted from their partial yields x_0, by the prediction equation $Z_0 = \bar{Z} + b\,(x_0 - \bar{x})$, there will be an additional variance due to this prediction(4) and the total variance of $\bar{Y}$ would become

$$\left[s_b'^2 \left\{ \frac{1}{m'} - \frac{1}{M'} \right\} + s_e^2 \left\{ \frac{1}{m'} + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma(x - \bar{x})^2} \right\} \right]$$

so that

$$t = \frac{\left\{ \bar{Y} - A \right\}}{\left[s_b'^2 \left\{ \frac{1}{m'} - \frac{1}{M'} \right\} + s_b'^2 \left\{ \frac{1}{m'} + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma(x - \bar{x})^2} \right\} \right]^{\frac{1}{2}}}$$

and hence

$$m' = \frac{(s_b'^2 + s_0^2)}{\left[\frac{A^2\,\bar{Y}^2}{(100)^2\,t^2} - \sigma_0^2 \left\{ \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma(x - \bar{x})^2} \right\} \right] + \frac{s_b'^2}{M'}} \quad (8)$$

Here, again, when M' = ∞, the equation becomes

$$m' = \frac{(s_b'^2 + s_a^2)}{\left[\frac{A^2K^2}{(100)^2 t^2} - s_e'^2 \left\{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma(x - \bar{x})^2}\right\}\right]}$$

Suitable values were given to x, M' and K in equation (8) and the corresponding values for m' were obtained.

In equation (6), the difference I - A, being the excess of the simple daughter average of the bull over the herd average A, is equated to the real superiority $\frac{AK}{100}$ of the bull (over the herd average). Actually, as I depends on the sample selected, the difference I - A does not always equal the true difference $\frac{AK}{100}$ but only varies round it. As such, using the value of the number of the daughters m' obtained from (7) (or (8)), the K per cent. superiority of the bull will be detected in only 50 per cent. of the cases. To find out the value of m' required for obtaining a significant result with a greater probability (greater than .50), the following procedure may be followed:

Let I, A, K and other notations have the same meaning as in the earlier case.

Let

$$t_1 = \frac{(X - A)}{\left[\left\{\frac{1}{m'} - \frac{1}{M'}\right\} \sigma_b'^2\right]^{\frac{1}{2}}}.$$

t_1 is then the Student - t (one tailed) with $m' - 1$ degrees of freedom.

The quantity

$$\frac{\left\{ (X - A) + \left\{ \frac{A K}{100} \right\} \right\}}{\left\{ \left\{ \frac{1}{m'} - \frac{1}{M'} \right\} \sigma_b'^2 \right\}^{\frac{1}{2}}} = t_2$$

follows the t - distribution with $m' - 1$ degrees of freedom.

If P denotes the probability of obtaining a significant result, then P is the probability that a value of t, with $(m' - 1)$ degrees of freedom, should exceed t_2. In the t - table, let p_2 denote the probability that a value lies outside the limits $\pm t_2$. P would then be $\left\{ 1 - \frac{p_2}{2} \right\}$. conversely, if the probability P is specified, the probability p_2 can be obtained $\left\{ = 2\,(1 - P) \right\}$ and the corresponding t_2 can be known from t - table.

Elimination of $(X - A)$ from the equations giving t_1 and

t_2 leads to

$$\left[(t_1 + t_2)^2 \cdot s_b'^2 \left\{ \frac{1}{m'} - \frac{1}{M'} \right\} \right]^{\frac{1}{2}} = \frac{A\,K}{100}$$

or $$\left[\frac{A^2\,K^2}{s_b'^2\,(100)^2\,(t_1 + t_2)^2} + \frac{1}{M'} \right] = \frac{1}{m'}$$

or $$m' = \frac{1}{\left\{ \dfrac{A^2\,K^2}{s_b'^2\,(100)^2\,(t_1 + t_2)^2} + \dfrac{1}{M'} \right\}} \qquad \text{———} (9)$$

when $M' = \infty$,

$$m' = \frac{1}{\left\{ \dfrac{A^2\,K^2}{(100)^2\,s_b'^2\,(t_1 + t_2)^2} \right\}} = \left\{ \frac{s_b'^2\,(100)^2\,(t_1 + t_2)^2}{A^2\,K^2} \right\}$$

If only partial records are used, the 't' values would be

$$t_1 = \frac{(\bar{Y} - A)}{\left[s_b'^2 \left(\frac{1}{m'} - \frac{1}{M'} \right) + s_e^2 \left\{ \frac{1}{m'} + \frac{1}{n'} + \frac{(x_0 - \bar{x})^2}{\Sigma (x - \bar{x})^2} \right\} \right]^{\frac{1}{2}}}$$

$$t_2 = \frac{\left\{ (Y - A) - \frac{AK}{100} \right\}}{\left[\left\{ s_b'^2 \left(\frac{1}{m'} - \frac{1}{M'} \right) + s_e^2 \left(\frac{1}{m'} + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma (x - \bar{x})^2} \right) \right\} \right]^{\frac{1}{2}}}$$

and hence

$$m' = \frac{s_b'^2 + s_e^2}{\left[\frac{A^2 K^2}{(100)^2 (t_1 + t_2)^2} - s_e^2 \left\{ \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma (x - \bar{x})^2} \right\} + \frac{s_b'^2}{M'} \right]} \qquad (10)$$

Here, again, $M' = \infty$ gives

$$m' = \frac{s_b'^2 + s_e^2}{\left[\frac{A^2 K^2}{(100)^2 (t_1 + t_2)^2} - s_e^2 \left\{ \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma (x - \bar{x})^2} \right\} \right]}.$$

Equations (10), and (9), can thus be used to get the values of m' for different probabilities P (of obtaining a significant result) at different values of M', K and x. For a given M', K and x, m' would naturally increase with P. For want of time the values of m' using (9) and (10) could not be computed.

It may be mentioned that to find out the number of cows required for establishing as significant a specified percentage superiority of a new treatment over a standard treatment, the procedure would be same as in the earlier case of bull, the only difference being that the variances in the "t" tests now would be twice the variances of those in the earlier case.

4. 6. Prediction of lactation yield when partial yield itself is estimated from a sample of daily yields.

As seen earlier (4), the variance of the predictor, giving the total lactation yield Z_0 of a cow from her observed partial yield x_0, is

$$\left\{ s_o^2 \left\{ 1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{\Sigma (x - \bar{x})^2} \right\} \right\}.$$

If x yield itself is estimated from a sample of yields on certain selected days of the initial period, the variance of the predictor would be increased to

$$\left\{ s_e^2 \left\{ 1 + \frac{1}{n} + \frac{(x_s - \bar{x})^2}{\Sigma (x - \bar{x})^2} \right\} + b^2 \ V(x_s) \right\}$$

where x_s is the estimated partial yield. Thus the increase in the variance of the predictor due to sampling is $b^2 \ V(x_s)$

and hence the loss of efficiency due to sampling is

$$\frac{b^2 V(x_s)}{\left\{ s_e^2 \left\{ 1 + \frac{1}{n} + \frac{(x_s - \bar{x})^2}{\Sigma (x - \bar{x})^2} \right\} + b^2 \; V(x_s) \right\}}$$

For a given group of m cows, the loss of efficiency in adopting sampling procedure is hence

$$\frac{\left\{ b^2 \; \frac{\overline{V(x_s)}}{m} \right\}}{\left\{ s_e^2 \left\{ \frac{1}{m} + \frac{1}{n} + \frac{(\bar{x}_s - \bar{\bar{x}})^2}{\Sigma (x - \bar{x})^2} \right\} + b^2 \frac{\overline{V(x_s)}}{m} \right\}}$$

where $\bar{x}_s$ is the average estimated partial yield for the m cows and $\overline{V(x_s)}$ is the average of the m variances $V(x_s)$ for the m cows. $V(x_s)$ depends upon the method of sampling employed.

Two methods of sampling, namely, simple random sampling and systematic sampling were studied.

1. <u>Simple Random Sampling.</u>

Let d_1, d_2 ———— d_n respectively denote the milk yields of a cow on the 1st, 2nd ———— nth days so that the total yield for the initial period of the first N days is $\sum_{i=1}^{N} d_i = x(N)$, say. Out of these N days, k days are selected at random and yields are observed on these k selected days. Let $x(k)$ be the total of the yields in these k days, i.e., $x(k) = \sum_{i=1}^{k} d_i$, summation being taken over the k selected days. Let $\bar{x}_{(N)} = \frac{x(N)}{N}$ and $\bar{x}_{(k)} = \frac{x(k)}{k}$. Then $x_s = N\,\bar{x}_{(k)}$ gives an unbiased estimate of the partial yield $x_{(N)}$, and its variance is

$$V(x_s) = N^2 \frac{N-k}{Nk} S^2 \qquad \text{where}$$

$S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (d_i - \bar{x}_{(N)})^2$ is the mean square between daily yields on the N days. For a given m cows the variance would be $N^2 \times \frac{N-k}{Nk} \times \frac{\bar{S}^2}{m}$ where $\bar{S}^2$ is the average of the mean square S^2 for the m cows.

2. Systematic Sampling.

Let, again, d_1 ------ d_N be the yields on the 1st, 2nd ------ N^{th} days so that the total yield for the initial period of N days is $\sum_{i=1}^{N} d_i = x_{(N)}$.

If the daily yields are recorded systematically at intervals of r days starting with a randomly chosen day, say day, with in the first interval $(1 \leq i \leq r)$, the yields will be d_i, $d_i + r$, $d_i + 2r$ ——— $d_1 + (k-1)r$, k being the size of the sample.

An unbiased estimate of the total yield $x_{(N)}$ for N days is provided by the sample total $x_s = r \sum_{l=1}^{k-1} d_{i+lr}$

The variance of this estimate is given by

$$\bar{S}^2_w(r) = \left\{ r \sum_{i=1}^{r} x_i'^2 - x_{(N)}^2 \right\}$$

For given m cows, their average yield for N days is similarly estimated and the variance of the estimate would be

$$\frac{\bar{S}^2_w(r)}{m} \quad \text{where} \quad \bar{S}^2_w(r) \quad \text{is the average of the}$$

$S^2_w(r)$ values for the m cows.

By comparing the two variances, namely, $\frac{N(N-k)}{k} S^2$ and $S^2_w(r)$, the efficiency of systematic sampling as compared to simple random sampling can be worked out for each of the cows in a herd.

Thus, if the average partial yield x of m cows is estimated, instead of being observed, from a sample of daily yields, the loss of efficiency is

$$\frac{\left\{ b^2 \frac{N(N-k)}{k} \quad \frac{\bar{S}^2}{m} \right\}}{\left\{ s_o^2 \left\{ \frac{1}{m} + \frac{1}{n} + \frac{(x_s - \bar{x})^2}{\Sigma(x-\bar{x})^2} \right\} + \left\{ b^2 \frac{N(N-k)}{k} \quad \frac{\bar{S}^2}{m} \right\} \right\}} \qquad (11)$$

if the sample is a simple random sample of k yields from among the N yields, and is

$$\frac{\left\{ b^2 \quad \frac{\bar{S}w^2(r)}{m} \right\}}{\left\{ s_o^2 \left\{ \frac{1}{m} + \frac{1}{n} + \frac{(x_s - \bar{x})^2}{\Sigma(x-\bar{x})^2} \right\} + \left\{ b^2 \quad \frac{\bar{S}w^2(r)}{m} \right\} \right\}} \qquad (12)$$

if the sampling procedure is a 'SystematicSampling' technique with the size of the sample k and the interval of recording r.

For one of the herds, namely, the Red Sindhi herd at Hosur, the percentage standard error of the estimate for the yield total $x_{(N)}$ for 112 days from a simple random sample of 16 days and that of the corresponding estimate from systematic sampling at weekly intervals of recording were worked out for each cow and the efficiency of the latter technique as compared with the former method was estimated. The loss of efficiency resulting from adopting either of the two systems of sampling was then worked out for different values of the group size m.

5. RESULTS.

1. Table IIa gives the values of the correlation coefficient r_{xz} between total lactation yield z and partial yield x for different periods of initial yield, the target being the period for which the square of the correlation (r^2_{xz}) is of the order of .80. The table also gives the values of the coefficient for initial yield upto the peak week. The percentages variation in Z explained by its linear regression on x, being the squares of the corresponding correlations, are also given in the table.

The values of the correlation coefficient r_{xy} between partial yield x and later yield y corresponding to the r_{xz} values in Table IIa are given in Table IIb; the squares of the correlations r_{xy} are also given in the table.

Table IIa.

Correlation r_{xz} *(together with its square - given in bracket) between total lactation yield and yield total upto peak week and week where the square of the correlation exceeds .80.

No.	Herd.	No. of cows.	r_{xz} for x upto peak week: Peak week No.	r_{xz} for x upto peak week: r_{xz}	r_{xz} for x upto 8 weeks.	12 weeks.	16 weeks.	20 weeks.	24 weeks.	28 weeks.
1.	Red Sindhi, Hosur.	34	13	.8772 (.76[illegible])	.6358 (.4041)	.8647 (.7477)	.8956 (.8021)			
2.	Red Sindhi, Bangalore.	40	8	.8594 (.7386)	.8594 (.7386)	.8574 (.7375)	.9065 (.8217)			
3.	Tharparkar, Patna (weaned)	65	8	.8439 (.7122)	.8439 (.7123)	.8810 (.7762)	.9189 (.8444)			
4.	Tharparkar, Patna (not weaned).	33	13	.8273 (.6852)	.8125 (.6601)	.8231 (.6775)	.8326 (.6932)	.8496 (.7218)	.8810 (.7744)	.9219 (.8499)
5.	Hariana, I.V.R.I.	33	4	.7060 (.4984)	.8021 (.6434)	.8557 (.7322)	.8994 (.8089)			
6.	Hariana, Hissar.	35	14	.7283 (.5304)	.6806 (.4632)	.7580 (.5746)	.7701 (.5930)	.8159 (.6657)	.8585 (.7370)	.9074 (.8234)
7.	Cost of milk production	31	11	.7751 (.6008)	.7313 (.5348)	.8048 (.6477)	.8583 (.7367)	.8950 (.8010)		

*The values are all positive.

Table IIb.

Correlation r_{xy} *(together with its square - given in bracket) between partial yield and later yield corresponding to correlation r_{xz} of Table IIa.

No.	Herd.	No. of cows.	r_{xy} for x up-to peak week. Peak week No.	r_{xy}	r_{xy} for x upto 8 weeks.	12 weeks.	16 weeks.	20 weeks.	24 weeks.	28 weeks.
1.	Red Sindhi, Hosur.	34	13	.6958 (.4841)	.7453 (.5553)	.7183 (.5159)	.7100 (.5041)			
2.	Red Sindhi, Bangalore.	40	8	.7621 (.5808)	.7621 (.5808)	.7394 (.5467)	.7121 (.5071)			
3.	Tharparkar, Patna(weaned).	65	8	.7671 (.5884)	.7671 (.5884)	.7783 (.6057)	.8021 (.6434)			
4.	Tharparkar, Patna(not weaned).	33	13	.6843 (.4683)	.7263 (.5275)	.6897 (.4757)	.6436 (.4142)	.6006 (.3608)	.5748 (.3304)	.5681 (.3227)
5.	Hariana, I.V.R.I.	36	4	.6397 (.4092)	.6990 (.4889)	.7304 (.5335)	.7428 (.5517)			
6.	Hariana, Hissar.	35	14	.4856 (.2358)	.5482 (.3005)	.5881 (.3459)	.5059 (.2559)	.5342 (.2854)	.5551 (.3081)	.5556 (.3087)
7.	Cost of milk production survey, Delhi.	31	11	.6042 (.3650)	.6083 (.3700)	.6343 (.4023)	.6473 (.4190)	.6276 (.3939)		

*The values are all positive.

It is seen from the table (II) that the correlation coefficients r_{xz} are fairly high. It is also observed that the values show a definitely increasing trend with an increase in the length of the initial period. On the other hand, the r_{xy} values are seen to be erratic in behaviour.

The correlation r_{xz} for x yield upto the peak week is seen to be of the order of +.70 in the case of the Hariana herd at I.V.R.I., of the order of +.75 for the cows under the cost of milk production survey, Delhi and the Tharparkar herd, Patna (not weaned) and of the order of + .85 for the remaining four herds.

For the four herds, namely, the Red Sindhi herd at Hosur, the Red Sindhi herd at Bangalore, the Tharparkar herd (weaned) and the Hariana herd at I.V.R.I., r_{xz} is seen to be of the order of + .80 in the case of 8-week initial yield, the corresponding figures in the case of 12-week yield and in that of 16-week yield being + .85 and + .90 respectively. Correspondingly, the percentages variation in the total lactation yield as explained by its linear regression on partial yield are of the order of 64, 72 and 80 for the 8-week, 12-week and 16-week initial yields respectively. It is thus seen that the desired level of r^2_{xz} equal to .80 is reached with 16-week partial yield in the case of these four herds. For the Cost of milk production survey, Delhi a 20-week period is seen to be sufficient while for the other two herds, namely, Tharparkar (not weaned) and Hariana Hissar, the period required for attaining the desired correlation of + .90 is observed to be as long as 28 weeks. It is further seen that in the case of the last three herds mentioned above, the values of r^2_{xz} for 16-week initial period are respectively of the order of .74, .69 and .69. Thus 16-week initial yield is seen to be generally a good choice for any herd.

2. The seven fitted equations $Z_0 = a + bx_0$, along with the standard errors of b, corresponding to the seven herds are given in Table III below.

TABLE III.

Prediction equation and standard error of b.

No.	Herd.	Equation.	S.E.(b) (in lb.)
1.	Red Sindhi, Hosur.	$Z_0 = 305.33 + 2.0023\, x_0$	± .1787
2.	Red Sindhi, Bangalore.	$Z_0 = 188.91 + 1.8959\, x_0$	± .1408
3.	Tharparkar (weaned).	$Z_0 = -727.00 + 2.3318\, x_0$	± .1375
4.	Tharparkar (not weaned).	$Z_0 = -84.40 + 1.4034\, x_0$	± .1003
5.	Hariana, I.V.R.I.	$Z_0 = -734.33 + 2.3274\, x_0$	± .1693
6.	Hariana, Hissar.	$Z_0 = -101.36 + 1.5479\, x_0$	± .1249
7.	Cost of milk production survey, Delhi.	$Z_0 = -60.21 + 1.6716\, x_0$	± .1503

In the prediction equations given in the table (III), x_0 denotes the milk yield (lb.) for 16 weeks in the case of Red Sindhi, Hosur, Red Sindhi, Bangalore, Tharparkar (weaned) and Hariana, I.V.R.I. herds, for 20 weeks in the case of the Delhi cows and for 28 weeks in the case of the remaining two herds.

From the values of b it is seen that for a pound increase in the yield in the earlier period, one may expect on an average an increase of 2 lb. in the lactation yield in the case of the first four herds mentioned above and of about 1.5 lb. in the other three herds. Percentage standard error of b is seen to be quite low, the maximum value being of the order of 9 percent in the case of the Delhi survey.

3. Average predicted yields Z_0 along with actual and percentage errors of the predictor for different initial yields x_0 and for groups of different sizes (m) are given in the tables in the Appendix. In each of these tables the given range of values for x_0 covers more than 80 per cent. of the observed initial yields.

It will be seen from the tables that, as expected, the percentage standard error decreases as m is increased. For change in m from 1 to 6 the percentage error is seen to be reduced to half. The reduction afterwards is seen to be not so rapid, there being not much difference in the values for m=20 and even m = ∞.

For $m=1$, the percentage values are seen to be quite high. In most of the cases the error is seen to be between 10 per cent. and 20 per cent. It is thus seen that the predicting power of early yield is poor for a single cow. The minimum value for the percentage error is seen to be found, as is obvious, for $m = \infty$; in most of the cases, the value is seen to lie between 2 per cent. and 8 per cent. Along with these tables are given the tables giving the fiducial limits (95 per cent.) for Z_0 in the limiting cases of a single cow ($m = 1$) and an extremely large number of cows ($m = \infty$). The fiducial bands so obtained are represented in Figures (2).

As a corollary to the results in the Appendix, the values of the group size (m) for which it may be possible to predict the average lactation yield of the (given) group with five per cent. error of prediction were worked out and the results are given in table IV.

Table IV.

Size of the group (m) of different levels of production(x_0) for which prediction of average lactation yield Z_0 can be [illegible] with 5 per cent. error.

No.			Level of production and corresponding group size.		(Mean) $\bar{x}$		
1.	Red Sindhi, Hosur.	x_0	1000	1500	1897.12	2000	2500
		m	X	22	9	8	6
2.	Red Sindhi, Bangalore.	x_0	800	1000	1583.37	2500	3500
		m	X	X	10	5	5
3.	Tharparkar, (weaned)	x_0	1500	2000	2326.29	3500	4000
		m	X	63	12	8	6
4.	Tharparkar (not weaned).	x_0	2000	2500	3656.03	4500	5000
		m	X	186	10	7	6
5.	Hariana, I.V.R.I.	x_0	1500	2000	2233.00	3000	3500
		m	X	27	16	10	9
6.	Hariana, Hissar.	x_0	500		783.2	1000	1200
		m	X		6	4	3
7.	Cost of milk production Survey, Delhi.	x_0	1000		1675.63	2000	2500
		m	X		59	31	23.

X : Percentage error more than 5 per cent. for any m.

It is seen from the table (IV) that the value of m decreases for an increase in the level of production x_0. For the mean values $\bar{x}$, the values of m for the seven herds are seen to be respectively 9, 10, 12, 10, 16, 6 and 59. Except for the high value in the case of the Delhi cows, the values are quite moderate. One might conclude that, in general, the average lactation yield can be predicted with a standard error of 5 per cent. for a highly selected group of cows of even a small size such as about six while for attaining the same level of precision for average cows, the group must be about a dozen.

4. The percentages variation in the total lactation yield Z as explained by the three regression curves, namely, the St. line $Z = a+bx$, the quadratic $Z = a+b_1x+b_2x^2$ and the cubic $Z = a+b_1x+b_2x^2+b_3x^3$, obtained in the case of both the 12-week and the 16-week partial yields x, are given in table V

Table V.

Percentage variation in lactation yield as explained by its regression - linear and curvilinear - on 12-week and 16-week portial yields.

No.	Herd.	Regression of lactation yield on					
		Partial yield for 12 weeks.			Partial yield for 16 weeks.		
		Linear.	Quadratic.	Cubic.	Linear	Quadratic.	Cubic.
1.	Red Sindhi, Hosur.	77.62	77.72	78.83	84.44	84.58	86.47
2.	Red Sindhi, Bangalore.	74.73	75.23	76.42	80.22	80.23	84.50
3.	Tharparkar (weaned).	78.78	78.96	81.44	82.17	83.58	84.54
4.	Cost of milk production survey, Delhi.	64.77	65.47	65.98	72.67	73.71	73.60

It is seen from the table (V) that there is no appreciable difference in the percentages variation (in Z) explained by linear, second and third order regressions for both 12-week and 16-week initial yields.

For the Red Sindhi herd at Bangalore, the difference in the percentages variation for the st. line and the cubic curve is never more than three, the other differences, namely, between the st. line and the quadratic and the quadratic and the cubic being, obviously, still less. Similar is the case with the

Tharparkar herd. In the other two herds the differences are somewhat more but even so they are never more than five per cent.

6. Table VI gives the number of daughters n' required to be sampled from among M' cows to establish as significant, at 5 per cent. level of significance, the breeding worth of a bull expected to be superior to the given herd average A by K (20 and 10) per cent. when observed lactation yields are used and also when the lactation yields are predicted from partial yields x_0.

Table [illegible].

Values of the number of daughters n' for different values of percentage superiority K of the bull, different partial yields x_0 and for different population values N' of cows.

1) P = 20 per cent.

No.	Herd.	N' = 100: Total records used.	N' = 100: Partial records used, $x_0 = \bar{x}$	$x_0 = \bar{x}+\sigma$	$x_0 = \bar{x}+2\sigma$	N' = 500: Total records used.	N' = 500: Partial records used, $x_0 = \bar{x}$	$x_0 = \bar{x}+\sigma$	$x_0 = \bar{x}+2\sigma$	N' = ∞: Total records used.	N' = ∞: Partial records used, $x_0 = \bar{x}$	$x_0 = \bar{x}+\sigma$	$x_0 = \bar{x}+2\sigma$
1.	Red Sindhi, Hosur.	8	9	9	10	8	9	10	11	8	10	10	11
2.	Red Sindhi, Bangalore.	8	10	10	11	9	11	11	12	9	11	11	12
3.	Tharparkar (weaned).	12	14	14	15	13	15	16	17	13	15	16	17
4.	Tharparkar (not weaned).	10	11	12	13	10	12	13	14	11	12	13	14
5.	Hariana (I. V. R. I.).	11	12	13	14	11	13	14	15	11	14	14	15
6.	Hariana (Hissar).	5	6	7	8	6	8	8	8	7	8	8	8

47.

Table VI.(Contd.)

ii) K = 10 per cent.

No.	HERD.	N' = 100 Total records used.	N' = 100 Partial records used. x_0 = $\bar{x}$	$\bar{x}+\sigma$	$\bar{x}+2\sigma$	N' = 500 Total records used.	N' = 500 Partial records used. x_0 = $\bar{x}$	$\bar{x}+\sigma$	$\bar{x}+2\sigma$	N' = ∞ Total records used.	N' = ∞ Partial records used. x_0 = $\bar{x}$	$\bar{x}+\sigma$	$\bar{x}+2\sigma$
1.	Red Sindhi, Hosur.	21	23	33	62	26	34	41	100	28	38	44	120
2.	Red Sindhi, Bangalore.	24	31	35	58	29	39	45	87	31	41	49	101
3.	Tharparkar (weaned).	33	41	44	60	43	55	62	100	47	61	69	120
4.	Tharparkar (not weaned).	28	35	41	67	34	45	53	119	36	49	59	150
5.	Hariana, I.V.R.I.	29	41	40	60	37	54	66	93	39	60	73	114
6.	Hariana, Hissar.	18	22	24	38	20	25	28	43	21	28	30	47

σ^2: Variation in x between animals.

It is seen from the table (VI) that in detecting 20(K) per cent. superiority of the bull, the number m' of daughters to be sampled does not increase very much when the true size N' of the population is increased while in detecting 10 per cent. superiority the increase in m' is quite appreciable. Also, in detecting the 20 per cent. superiority, the increase in the value of m' on account of prediction is not much. For detecting 10 per cent. superiority, however, many more cows are required if only partial records are used instead of the total ones. It is further seen that for cows with superior level of initial production (x_0) the number of cows to be sampled is more than in the case of cows with average level of initial production $\bar{x}$, the difference being more pronounced for K = 10 than for K = 20.

6. Table VIIa gives the percentage standard errors of estimates of partial yield $x_{(n)}$ for 112 days under two systems of sampling, namely, simple random sampling and systematic sampling where the size of the sample, in both the cases, is sixteen (days) and the interval of recording in the latter case is seven days. The table gives the average percentage relative efficiency of systematic sampling with respect to simple random sampling of daily yields.

Table VII_b gives the percentage loss of efficiency due to sampling of the daily yields of the initial period in the case of both the methods of sampling for different values of the group size m and partial yields x_a. The results pertain to the Red Sindhi herd at Hosur.

Table VIIc.

Percentage standard errors p_1 and p_2 of systematic and simple random sampling estimates of partial yield, and average percentage relative efficiency P of systematic sampling with respect to simple random sampling.

Herd	No. of cows.	Interval of recording.	p_1	p_2	P
Red Sindhi, Hosur.	34	one week	8.62	4.65	849.

Table VII(b).

Percentage loss of efficiency due to systematic and simple random sampling of yields of the initial period for different levels of initial production (x_0) and group size (m).

	Simple Random Sampling.				Systematic Sampling.			
	m				m			
	1	5	10	20	1	5	10	20
1000	9.2	6.7	5.1	3.4	3.1	2.8	1.7	1.1
1500	9.7	8.4	7.1	5.6	3.3	2.8	2.4	1.8
1897.12	9.8	8.9	8.0	6.6	3.4	3.0	2.7	2.3
2000	9.8	8.8	7.9	6.5	3.4	3.0	2.6	2.1
2500	9.5	7.8	6.3	4.6	2.8	2.6	2.1	1.5

It is seen from table VII(a) that systematic sampling is more than eight times as efficient as simple random sampling.

It is seen from table VII that, as expected, loss of efficiency due to sampling decreases as m is increased. Also, for both the systems of sampling, the maximum loss is found, for all values of m, at th mean value $\bar{x}$ of partial yield. For a group of more than five, loss of efficiency in the case of systematic sampling is not more than three per cent. for any level of initial production.

6. DISCUSSION AND CONCLUSIONS.

1. The results on correlation show that the coefficient of correlation (r_{xz}) between lactation yield (z) and partial yield (x) is fairly high, in most of the cases, even for an initial period as short as eight weeks. There is a steady increase in the value as the period of initial yield is increased, and the level of +.90 for the correlation is reached in sixteen weeks in the case of Red Sindhi, Hosur, Red Sindhi, Bangalore, Tharparkar (weaned), Patna and Hariana, I.V.R.I. herds. Initial yield for sixteen weeks thus seems to be generally satisfactory as a basis for prediction of total lactation yield in well managed herds. In the case of the data on cows recorded in the Cost of milk production survey, Delhi, the correlation of +.90 is reached for initial yield upto twenty weeks while in the remaining two herds, a period of twenty eight weeks is required to attain the value.

2. The results indicate that initial yield upto the peak week does not serve as a good predictor of the total lactation yield. In some of the cases, half of the total variation remained unexplained by the yield upto the peak.

3. The results on correlation are, in general, in agreement with results in some of the earlier studies. Some workers have used 70 days' yield, with a correlation coefficient of the order of +.80 with total lactation yield, as a guide to prediction. In the present study, except for the Hariana herd at Hissar, 70 days' yield gave a correlation of more than +.80 . 100 days' yield is considered good by some workers. Here, again, the present study gave a correlation of more than +.85 in all the herds except the Hissar herd where the correlation was +.75 .

As pointed out earlier, however, it does not seem to be safe to base any reliance on prediction based on a correlation of less than +.90 since otherwise a large fraction of the variation in the total yield could remain unexplained.

4. The correlation r_{xz} between partial yield and total yield is subject to the criticism that it contains an element of spurious correlation due to the early yield being itself a part of the total yield. To obviate this the correlation r_{xy} between the partial yield x and the later yield y were worked out. However, the results show that the correlation values do

not increase consistently with an increase in the length of the initial period. It cannot hence be made the basis of deciding the stage at which early yield could be taken as a reliable guide for predicting the total lactation yield. It is known that in the regression of total yield on partial yield as well as in that of later yield on partial yield, the 'least squares' estimates of the regression coefficients are obtained by minimising the same quantity, the total yield being the sum of the partial yield and the later yield. The regression coefficients in the two differ just by unity. Also, since the residual variances are same, tests of significance of the regression coefficients would be identical. The problem of direct practical interest, moreover, is to predict the total lactation yield, and not the later yield, and as such it would be preferable to use the regression of total yield on partial yield directly to deriving it from the regression of later yield on the partial yield.

6. Correlation between total yield and yield total after the peak week, and other later yields were also worked out. Later yield after the peak week gave a correlation of more than

+.95 for all the herds. In addition, it was found that for identical periods, the total yield had greater correlation with later yield than with initial yield. However, correlation between later and total yields is not of wide practical importance except in the case of short term nutritional experiments in dairy animals which are traditionally conducted on animals which have attained their peak yield during the lactation.

6. The results on standard error of prediction show that the power of prediction is not high in the case of a single cows' yield, the standard error of predictor varying from about ten per cent, to thirty per cent. However, the average yield of a group of even a few cows, such as ten, is predicted fairly accurately, the standard error of prediction varying from about five to ten per cent. It is observed that, in general, relatively more accurate prediction is possible for cows with higher level of production than for cows with lesser yield. In fact, average lactation yield can be predicted with a standard error of five per cent for a highly selected group of cows of even a small size such as six while for attaining the same level of precision for average

cows, the group must be about a dozen.

7. The results regarding the percentage variation in total yield as explained by its regression - both linear and curvilinear - on partial yield clearly show that only very slight improvement is achieved, in all the four herds examined, by using curvilinear regression instead of linear regression. It is thus seen that it is not worthwhile to employ curvilinear regression.

8. We may consider the utility of prediction of lactation yield on the basis of early yield in breeding programmes.

In progeny testing such prediction does not help very much. In Indian herds, because of late maturity, by the time the milk records of daughters become available for comparison more than five years will have elapsed and as such a saving of about five to six months in the time required for testing as a result of basing the estimate of lactation yield of daughters on the early yield against loss of efficiency due to prediction is not worth aiming at especially in the case of low yielding cows.

In the case of heifers, however, it is highly profitable to have a precise method of predicting lactation yield in deciding on the male calves to be retained for possible breeding use on the basis of dams' lactation performance because, in this case, in the absence of such prediction, it would be necessary to retain all the male calves till the completion of the lactation of the dams and this, in turn, would mean a heavy expenditure on the maintenance of unwanted male calves.

The results on the number of daughters required to prove a Sire show that about ten daughters are required to be sampled to establish, at 5 per cent. level of significance, the merit of a bull twenty per cent. superior to the herd average. If the test of significance is to be based on only partial records of the daughters, the number would be increased by two. On the other hand, for establishing the significance of ten per cent. superiority, about thirty five daughters are needed and, in addition, if only partial records of the daughters are used, the number would be increased by a large extent especially when the level of initial production x deviates materially from the mean value $\bar{x}$ for the set of cows whose records are used for formulating the prediction equation.

9. Prediction can be used to advantage in shortening the period for carrying out experiments of certain types on dairy cattle. Here if it could be assumed that the level of production has no interaction with the treatment effects, one might take per treatment a small number - about five to ten - of superior cows (as determined by previous lactation yields) since in this case their average total lactation yield is predicted with only about five per cent. error.

10. It is clear from the results on 'sampling of milk yields of initial period' that the estimate of the partial yield x based on systematic sampling with weekly intervals of recording is more than eight times as efficient as is simple random sampling with the same number of records apart from being more convenient in practice. In this method of systematic sampling, partial yield of even a single cow is estimated with a precision corresponding to a standard error of the order of three per cent.

The percentage loss of efficiency because of estimation (of partial yield) is fairly low in the case of systematic sampling, being less than four per cent. in the case of a single cow's yield, and of the order of two per cent. for a group of ten. Thus it is seen that there is not much loss in precision in estimating the partial yield from a systematic sample of yields recorded at weekly intervals, and using this estimate in the prediction equation to get the final lactation yield.

11. For this Red Sindhi herd, Hosur, correlation was worked out between total lactation yield and estimated yield based on systematic sampling at eight weekly intervals throughout the lactation. Out of the twenty samples examined, the correlation coefficient was above +.80 in the case of fifteen, thirteen of which giving a correlation of more than +.90 . Thus the method is seen to be attractive if the milk recording can be spread through-out the lactation.

For want of time, the study on 'sampling' could not include other designs of sampling, such as systematic sampling at bi-weekly, monthly and other intervals, stratified random sampling etc., and also could not be made to cover all the herds.

SUMMARY.

Investigation, using data from seven Indian herds, has been carried out on the problem of predicting lactation yield of a cow from her partial yield.

1. Correlation coefficients were found between lactation yields and partial yields and also between partial yields and later yields. In the four herds, namely, Red Sindhi, Hosur, Red Sindhi, Bangalore, Tharparkar (weaned) and Hariana, I.V.R.I., sixteen weeks' yield, with a correlation of the order of +.90 with lactation yield, was found to be a good predictor. This initial period covered about 50 per cent. of the lactation yield and about 35 per cent. of the lactation length. In the case of cost of milk production survey, Delhi, an initial period of twenty weeks, covering about 60 per cent. of the lactation yield and about 50 per cent. of the lactation length, was found sufficient to attain the desired correlation of +.90. For the remaining two herds of Hariana (Hissar) and Tharparkar (not weaned) cows, the required period was as long as twenty eight weeks; this period covered about 60 per cent. of the lactation length and about the same (60) percentage of the lactation yield.

2. It was observed that the power of prediction was poor in the case of a single cow's yield, the standard error of the predictor varying from about 10 per cent. to thirty per cent. However, prediction could be made fairly accurately in the case of average yield of a given group of even a few cows such as ten. It was further observed that the percentage standard error of prediction was much lower for high yielding cows than for low yielding ones. Prediction line and its 95 per cent. fiducial bands have also been graphically represented.

3. Use of curvilinear regression was seen to be of only very slight help in improving the prediction. On the other hand, the proportion of variation explained by the linear prediction formula could be increased more by extending the period of initial yield.

4. About ten daughters were found to be necessary to establish the significance of the merit of a bull which is twenty per cent. superior to a given herd average. The use of partial records of the daughters would increase this number to about a dozen. For establishing ten per cent. superiority, however, as

many as thirty five daughters were found necessary, and this number would be increased by a large extent in the case of partial records.

6. The problem of predicting lactation yield from partial yield which itself is estimated from a sample of yields on randomly selected days of the initial period was also investigated. Systematic sampling at weekly intervals and equivalent simple random sampling were considered. The former procedure was found to be more than eight times as efficient as is the latter technique in estimating the partial yield. It was found, in addition, that loss of efficiency because of the sampling of partial yields was less than four per cent. in the case of a single cow's yield while for a group of ten the loss was only of the order of two per cent., the sampling technique being the 'Systematic sampling'. It was thus found that, at first, partial yield could be estimated from an appropriate sample and this estimated yield could then be used, with reasonable safety, in the prediction equation to arrive at, finally, the lactation yield.

REFERENCES.

1. Amble V.N., Rajagopalan, M. and Gill, B.S.(1953): 'Sampling of milk records for estimating lactation yield of cows.' -Indian J.Vet.Sci. and Anim.Husb. No.4, Vol.23, 13, 1953.

2. Bussort, K.(1957): 'Part yields of heifers in relation to the shape of the lactation curve and their value in estimating genetic merit.' -Anim. Breed. Abstr. No.619, Vol.25, 141, 1957.

3. Cannon, C.Y., Frye, J.B. and Sims, J.A.(1942): 'Predicting 305- day yields from short-time records.' -J. Dairy Sci.Vol.25, 991, 1942.

4. Gaines, W.L.(1927): 'The Deferred Short-time Test as a Measure of the Performance of Dairy cows.' -J. Agric. Research Vol.35, 237, 1927.

5. Gooch, M.(1935): 'An analysis of the time change in milk production in individual lactations' -J. Agric. Sci. Vol.25, 72, 1935.

6. Kartha, K.P.R.(1934): 'Prediction of milk yields.' -Indian J.Vet. Sci. and Anim. Husb. Vol.4, 218, 1934.

7. Mahadevan, P.(1951): 'The effect of Environment and Heredity of Lactation.' -J.Agric. Sci.Vol.41, 89, 1951.

8. McKellip, I and Seath, D. (1941):	'A comparison of different methods of calculating yearly milk and butterfat records.' -J. Dairy Sci. Vol.24, 181, 1941.
9. Mood, A.M. (1950)	- Introduction to the theory of Statistics.
10. Parou, C. (1949):	'Investigations on the possibility of determining the lactation yield of a milch cow from a short, initial lactation period.' -Anim. Breed. Abstr. No.83, Vol.18, 33, 1950.
11. Rendel et.al(1957):	'The Inheritance of Milk Production Characteristics.' -J.Agric.Sci.Vol.48,426, 1957.
12. Bognoni,G and Brambilla, E.(1957):	'Is it possible to simplify milk recording by limiting the time of collecting the production data?' -Anim.Breed.Abstr.No.1,Vol.26,38,1958.
13. Sanders,H.G.(1930):	'The analysis of the Lactation curve into Maximum Yield and Persistency.' -J.Agric.Sci.Vol.20,145, 1930.
14. Saxena,P.N. and Kumar,S. (1959):	'On Prediction of Total Yields.' - Unpublished.
15. Tuff, Professor (1928):	'External factors affecting milk yield.' -World Dairy Congress, 1928.

APPENDIX.

Table 1.1.

Average predicted yields ($Z_o = 335.36 + 2.0023\, x_o$) along with actual and percentage errors of prediction for different partial yields (x_o) and for groups of different sizes (m).

(Red Sindhi Herd, Hosur.)

Partial yields x_o (lb.)	1000		1500		Mean 1897.12		2000		2500	
Predicted yields Z_o (lb.)	2337.66		3338.81		4133.96		4339.96		5341.11	
m	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.
1	504.60	21.9	573.91	17.0	572.54	13.7	572.83	13.0	582.66	10.8
5	309.54	12.9	279.42	8.2	270.23	6.5	270.90	6.2	290.90	5.4
10	250.53	10.8	216.01	6.3	203.00	4.8	203.82	4.6	229.74	4.3
20	256.64	9.4	174.09	5.1	159.02	3.8	160.07	3.6	191.59	3.6
∞	187.07	7.8	119.95	3.5	96.77	2.3	93.50	2.2	144.70	2.7

Table 1.2.

Average predicted yields (Z_o= 185.91 + 1.8959 x_o) along with actual and percentage errors of prediction for different partial yields (x_o) and for groups of different sizes (n).

(Red Sindhi Herd, Bangalore.)

Partial yields x_o(lb.)	800		1000		Mean 1587.37		2500		3500	
Predicted yield Z_o(lb.)	1702.65		2081.71		3187.63		4925.41		6821.53	
n	S.E(lb.)	%S.E.	S.E(lb.)	%S.E.	S.E(lb.)	%S.E.	S.E(lb.)	%S.E.	S.E(lb.)	%S.E.
1	464.03	27.2	458.16	22.0	450.74	14.1	468.85	9.5	525.33	7.7
5	238.24	14.0	226.59	10.9	211.18	0.6	247.49	6.0	342.65	5.0
10	192.20	11.3	177.54	8.5	157.40	4.9	203.54	4.1	312.39	4.6
20	104.40	9.7	147.00	7.1	121.02	3.8	177.54	3.6	233.14	4.3
∞	130.84	7.7	108.17	5.2	70.39	3.2	147.00	3.0	273.83	4.1

Table 1.3.

Average predicted yields ($y_0 = 727.00 + 2.3018\ x_0$)
along with actual and percentage errors of prediction
for different partial yields x_0 and for groups of
different sizes (n).

(Thumparkar Nord (second), Patna.)

Partial yields x_0 (lb.)	1500		2000		Mean 2626.23		3500		4000	
Predicted yields y_0 (lb.)	2815.7		3906.6		5948.13		7530.3		8720.2	
n	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.
1	986.04	35.0	977.12	24.4	971.42	10.3	975.21	12.9	982.88	11.3
5	470.31	17.0	453.64	11.5	447.41	7.5	455.58	6.0	471.73	5.4
10	368.57	13.1	344.00	8.6	327.50	5.5	338.50	4.5	330.05	4.1
20	238.35	10.6	268.09	6.7	240.50	4.1	231.10	3.5	288.39	3.3
∞	207.14	7.3	160.38	4.0	119.60	2.0	147.20	1.9	191.68	2.2

Table 1.4.

Average predicted yields ($Z_o = -84.40 + 1.4084\, x_o$) along with actual and percentage errors of prediction for different partial yields (x_o) and for groups of different sizes (m).

(Tharparkar Herd (not weaned), Patna.)

Partial yields x_o (lb.)	2000		2500		Mean 3[illegible].03		4500		5000	
Predicted yields Z_o (lb.)	2732.40		3436.60		5034.75		6253.40		6957.60	
m	S.E.(lb)	% S.E.	S.E.(lb)	% S.E.	S.E.(lb)	% S.E.	S.E.(lb)	% S.E.	S.E.(lb)	% S.E.
1	730.93	27.0	720.25	21.2	719.94	14.2	724.92	11.6	732.50	10.5
5	378.83	13.8	357.48	10.4	338.07	6.7	348.55	5.6	364.05	5.2
10	302.49	11.1	278.01	8.1	252.57	5.0	266.43	4.3	283.42	4.1
20	257.30	9.4	228.12	6.6	196.31	3.9	213.85	3.4	233.20	3.4
∞	202.43	7.4	133.67	4.8	116.28	2.3	143.12	2.3	177.57	2.5

Table 1.5.

Average predicted yields ($Z_0 = -734.32 + 2.3274\ x_0$) along with actual and percentage errors of prediction for different partial yields (x_0) and for groups of different sizes (n).

(Hariana Herd, I.V.R.I.)

Partial yields x_0 (lb)	1500		2000		Mean 2233		3000		3500	
Predicted yields Z_0 (lb.)	2720.78		3890.43		4432.76		6217.28		7381.58	
n	S.E.(lb.)	%S.E.	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.	S.E.(lb)	%S.E.
1	769.07	28.2	756.77	19.4	705.33	17.0	757.21	16.0	735.00	10.8
5	383.84	14.1	358.56	9.2	356.61	8.0	359.48	7.1	431.56	5.9
10	303.01	11.1	270.27	6.9	266.34	6.0	271.49	5.3	335.13	4.9
20	253.12	9.3	212.81	5.5	207.79	4.7	214.33	4.2	324.90	4.4
∞	190.53	7.0	132.40	3.4	124.18	2.8	134.87	2.7	278.94	3.8

Table 1.6.

Average predicted yields ($Z_0 = -101.35 + 1.5479\ x_0$) along with actual and percentage errors of prediction for different partial yields (x_0) and for groups of different sizes (n).

(Hariana Herd, Hissar.)

Partial yields x_0 (lb.)	500		782.2		1000		1200	
Predicted yields Z_0 (lb.)	672.60		1109.42		1446.55		1756.13	
n	S.E.(lb.)	% S.E.	S.E.(lb.)	%S.E.	S.E.(lb.)	%S.E.	S.E.(lb.)	%S.E.
1	134.78	20.0	126.03	11.4	128.95	8.9	136.42	7.8
5	76.22	11.3	69.42	6.3	65.35	4.5	79.08	4.5
10	65.30	9.7	44.66	4.0	52.21	3.6	68.60	3.9
20	59.09	8.8	34.83	3.1	44.20	3.0	62.90	3.6
∞	52.15	7.7	21.77	2.0	34.37	2.4	56.25	3.2

Table 1.7.

Average predicted yields ($\hat{y}_0 = -30.2 + 1.0715\, x_0$) along with actual and percentage errors of prediction for different partial yields (x_0) and for groups of different sizes (n).

(Cost of Milk Production Survey, Delhi.)

Partial yields x_0 (lb.)	1000		1376.68		2000		2500	
Predicted yields $\hat{y}_0$ (lb.)	1591.29		2720.03		3202.79		4003.64	
n	S.E.(lb.)	%S.E.	S.E.(lb.)	%S.E.	S.E.(lb.)	%S.E.	S.E.(lb.)	%S.E.
1	631.66	39.7	623.44	22.9	625.35	19.2	630.70	15.6
5	312.67	19.6	295.72	10.9	299.71	9.2	320.63	7.8
10	245.10	15.4	223.15	8.2	223.42	7.0	255.24	6.2
20	203.19	12.8	175.99	6.5	192.02	6.0	215.23	5.2
∞	149.83	9.4	110.21	4.0	120.61	3.7	165.32	4.0

111.

Table 2. 1.

95 per cent. fiducial limits (U_1, U_2) for predicted yields Z_0 (= 385.36 + 2.0023 x_0) in the limiting cases of a single cow (m = 1) and an extremely large number of cows (m = ∞).

(Red Sindhi Herd, Hosur.)

Partial yield. x_0	Predicted yield Z_0	Fiducial limits.			
		m = 1		m = ∞	
		U_1	U_2	U_1	U_2
500	1386.51	119.86	2653.16	841.49	1931.53
1000	2387.66	1176.07	3599.25	2006.41	2768.91
1897.12	4183.96	3017.12	5350.80	3986.74	4381.18
2500	5391.11	4203.85	6578.37	5096.21	5686.01
3500	7393.41	6089.07	8697.75	6778.03	8008.79

Table 2.2.

95 per cent. fiducial limits (U_1, U_2) for predicted yields Z_0(= 165.91 + 1.8359 x_0) in the limiting cases of a single cow (n = 1) and an extremely large number of cows (n = ∞).

(Red Sindhi Herd, Bangalore.)

Partial yield x_0	Predicted yield Z_0	Fiducial limits. n = 1 U_1	n = 1 U_2	n = 2 U_1	n = 2 U_2
800	1702.55	763.89	2642.21	1437.60	1967.50
1000	2031.71	1153.94	3003.48	1862.66	2300.75
1583.37	3187.60	2274.91	4100.41	3045.13	3330.20
2500	4925.41	3973.93	5874.83	4627.73	5223.08
2500	6321.56	5757.77	7885.35	6256.87	7336.25

Table 2.3.

95 per cent fiducial limits (U_1, U_2) for predicted yields Z_0 (= 72700 + 2.3318 x_0) in the limiting cases of a single cow ($n = 1$) and an extremely large number of cows ($n = \infty$).

(Tharparkar Herd (weaned), Patna.)

Partial yield x_0	Predicted yield Z_0	Fiducial limits. $n = 1$ U_1	$n = 1$ U_2	$n = \infty$ U_1	$n = \infty$ U_2
1500	2315.7	843.62	4787.78	2401.42	3222.98
2000	3336.6	2042.36	5250.84	3677.84	4318.36
2826.23	5048.13	4005.29	7890.97	5700.13	6187.13
3500	7539.3	6588.88	9489.72	7244.90	7833.70
4000	8549.3	6754.44	10885.96	8337.04	9103.36

Table 2.4.

95 per cent. fiducial limits (U_1, U_2) for predicted yields Z_0 (= $-84.40 + 1.4084\, x_0$) in the limiting cases of a single cow (m = 1) and an extremely large number of cows (m = ∞).

(Tharpurkar land (not manured), Patna.)

Partial yield x	Predicted yield z	Fiducial limits. m = 1 U_1	m = 1 U_2	m = ∞ U_1	m = ∞ U_2
2000	2732.40	1233.11	4231.69	2321.57	3143.23
2500	3436.60	1956.95	4916.25	3164.57	3708.73
3656.03	5064.75	3603.99	6525.51	4920.81	5208.69
4500	6253.40	4782.53	7724.26	5963.01	6543.79
5000	6957.60	5471.36	8443.84	6597.31	7317.89

Table 2.6.

95 per cent. fiducial limits (U_1, U_2) for predicted yields Z_0 (= -704.32 + 2.3274 x_0) in the limiting cases of a single cow (m = 1) and an extremely large number of cows (m = ∞).

(Hariana Herd, I.V.R.I.)

Partial yield x_0	Predicted yield Z_0	Fiducial limits.			
		m = 1		m = ∞	
		U_1	U_2	U_1	U_2
1500	[illegible]	1162.40	4201.07	2380.24	3114.82
2000	3950.48	2851.21	5429.75	3621.18	4160.78
2233	4432.70	[illegible]	6960.20	[illegible]	4685.34
3000	6217.88	4650.97	7704.79	5019.91	6315.35
3500	7381.58	5783.33	8309.83	6314.22	7843.34

Table 2.9.

95 per cent. fiducial limits (U_1, U_2) for predicted yields Z_0 (= -101.35 + 1.5479 x_0) in the limiting cases of a single cow (m = 1) and an extremely large number of cows (m = ∞).

(Hariana Herd, Hissar.)

Partial yield. x_0	Predicted yield Z_0	Fiducial limits.			
		m = 1		m = ∞	
		U_1	U_2	U_1	U_2
500	672.60	333.19	947.01	536.42	778.70
732.2	1103.42	852.80	13[illegible]3.04	1005.09	1153.73
1000	1446.55	1134.00	1760.09	1378.57	1516.53
1200	1756.13	1478.33	2033.83	1641.[illegible]0	1870.6[illegible]

Table [illegible].7.

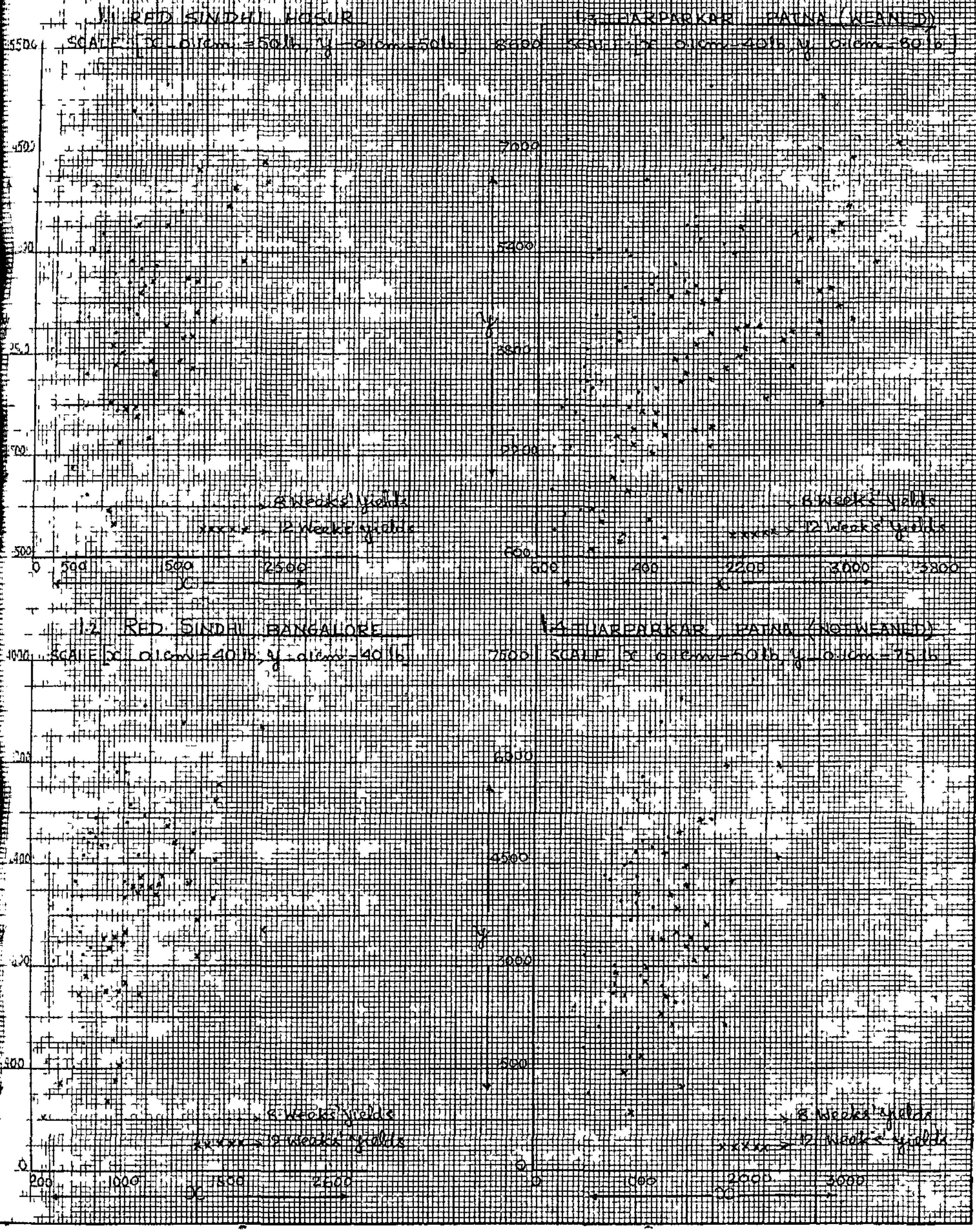
95 per cent. fiducial limits (U_1, U_2) for predicted yields z (= -80.2 + 1.6715 x_0) in the limiting cases of a single cow ($n = 1$) and an extremely large number of cows ($n = \infty$).

(Cost of milk production survey, Delhi.)

Partial yield. x_0	Predicted yield. z_0	Fiducial limits.			
		$n = 1$		$n = \infty$	
		U_1	U_2	U_1	U_2
1000	1591.29	290.54	2883.03	1204.83	1897.75
1675.63	2720.69	1145.75	3[illegible].62	2[illegible].31	2946.07
2000	3262.79	1903.05	4541.63	3013.35	3500.23
[illegible]	4[illegible].61	2708.35	6330.73	37[illegible].44	4437.03

FIGURES 1.

SCATTER DIAGRAMS OF INITIAL YIELDS (x) AND LATER YIELDS (y) FOR INITIAL PERIODS OF EIGHT AND TWELVE WEEKS



FIGURES 1.

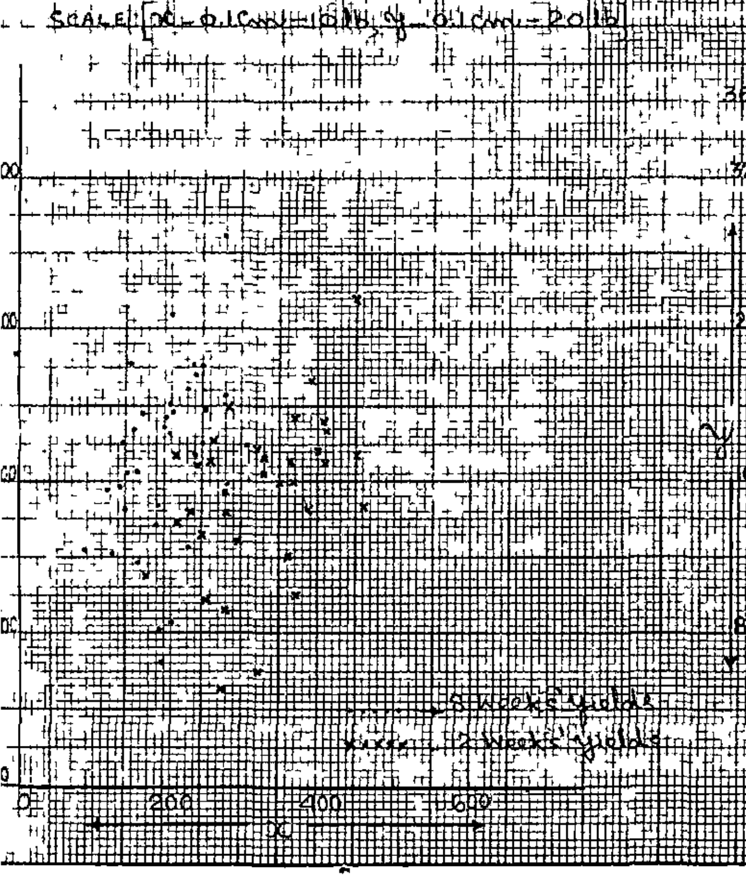
SCATTER DIAGRAMS OF INITIAL YIELDS (x) AND LATER YIELDS (y) FOR INITIAL PERIODS OF 8 AND 12 WEEKS

15 HARIANA, I.V.R.

SCALE [x 0.1cm = 40lb, y 0.1cm = 50lb]

16 HARIANA, HISSAR

SCALE [x 0.1cm = 10lb, y 0.1cm = 20lb]



17 COST OF MILK PRODUCTION SURVEY, DELHI

SCALE [x 0.1cm = 25lb, y 0.1cm = 40lb]

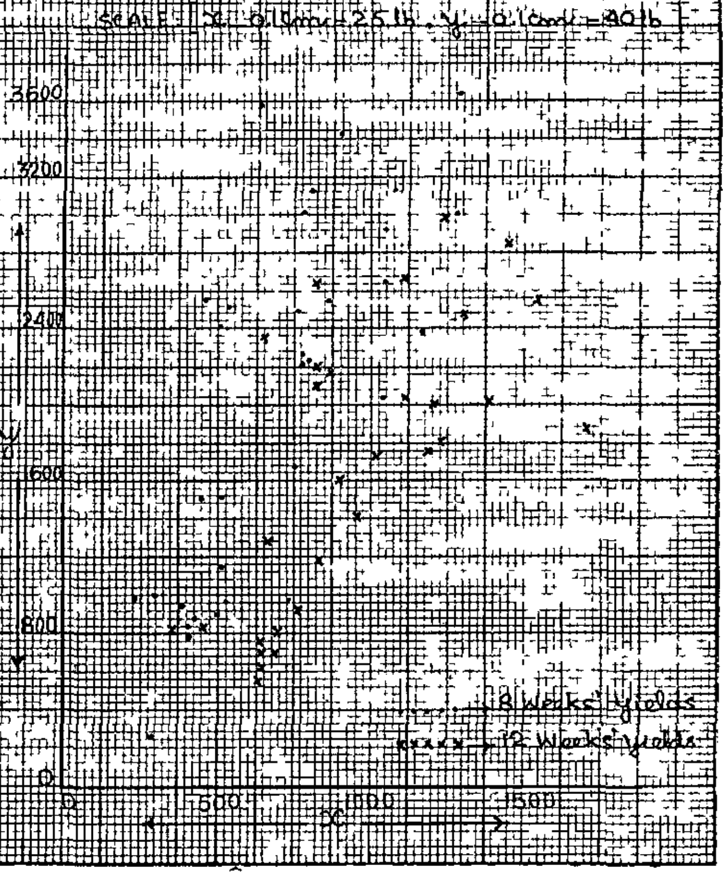
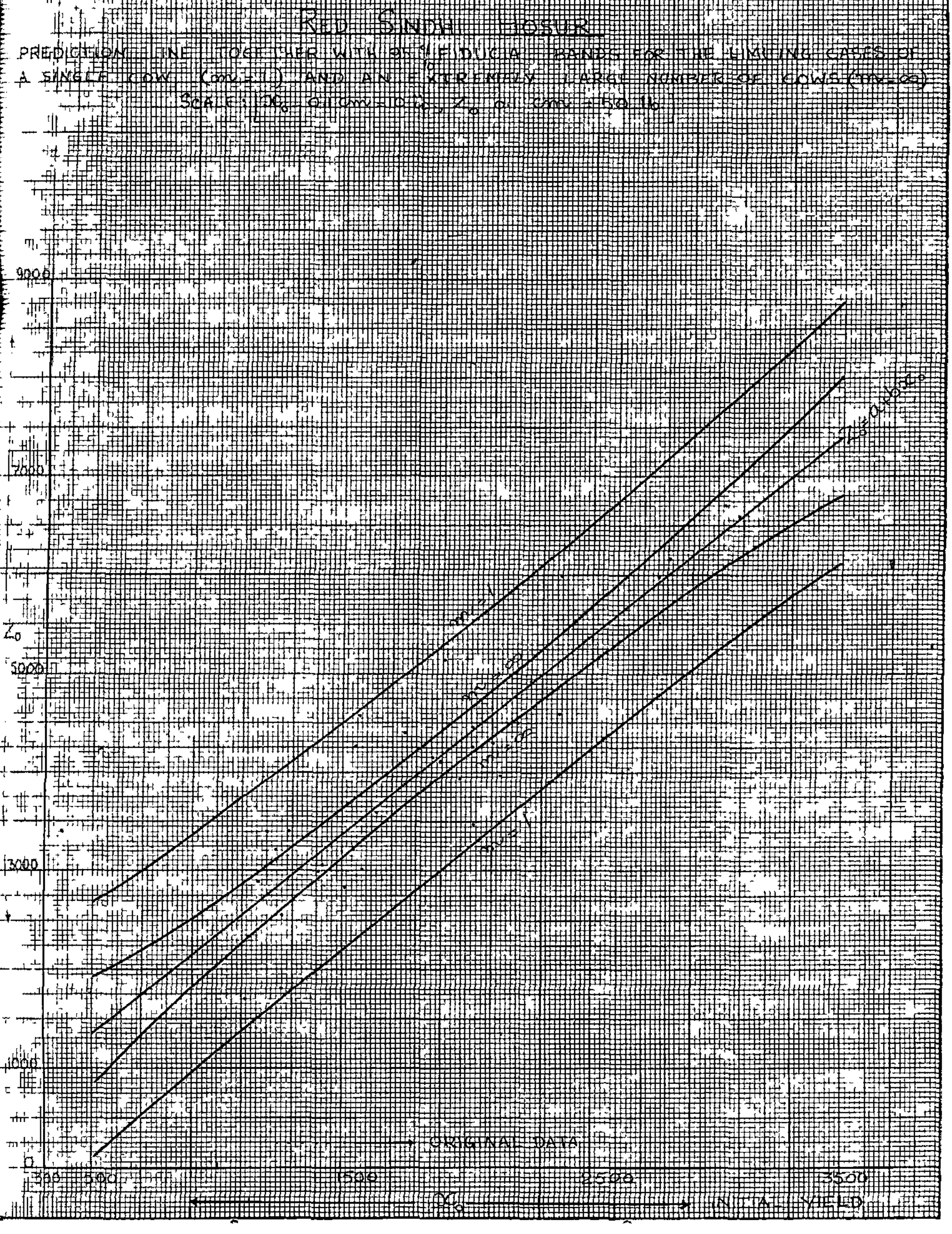


FIGURE 2.1.



RED SINDHI HOSUR
PREDICTION LINE TOGETHER WITH 95% FIDUCIAL BANDS FOR THE LIMITING CASES OF A SINGLE COW (m=1) AND AN EXTREMELY LARGE NUMBER OF COWS (m=∞)
9000
7000
Z_0
5000
3000
1000
0
300
500
1500
2500
3500
ORIGINAL DATA
x_0
N.T.A. YIELD
m=1
m=∞
m=∞
m=1

FIGURE 2.2

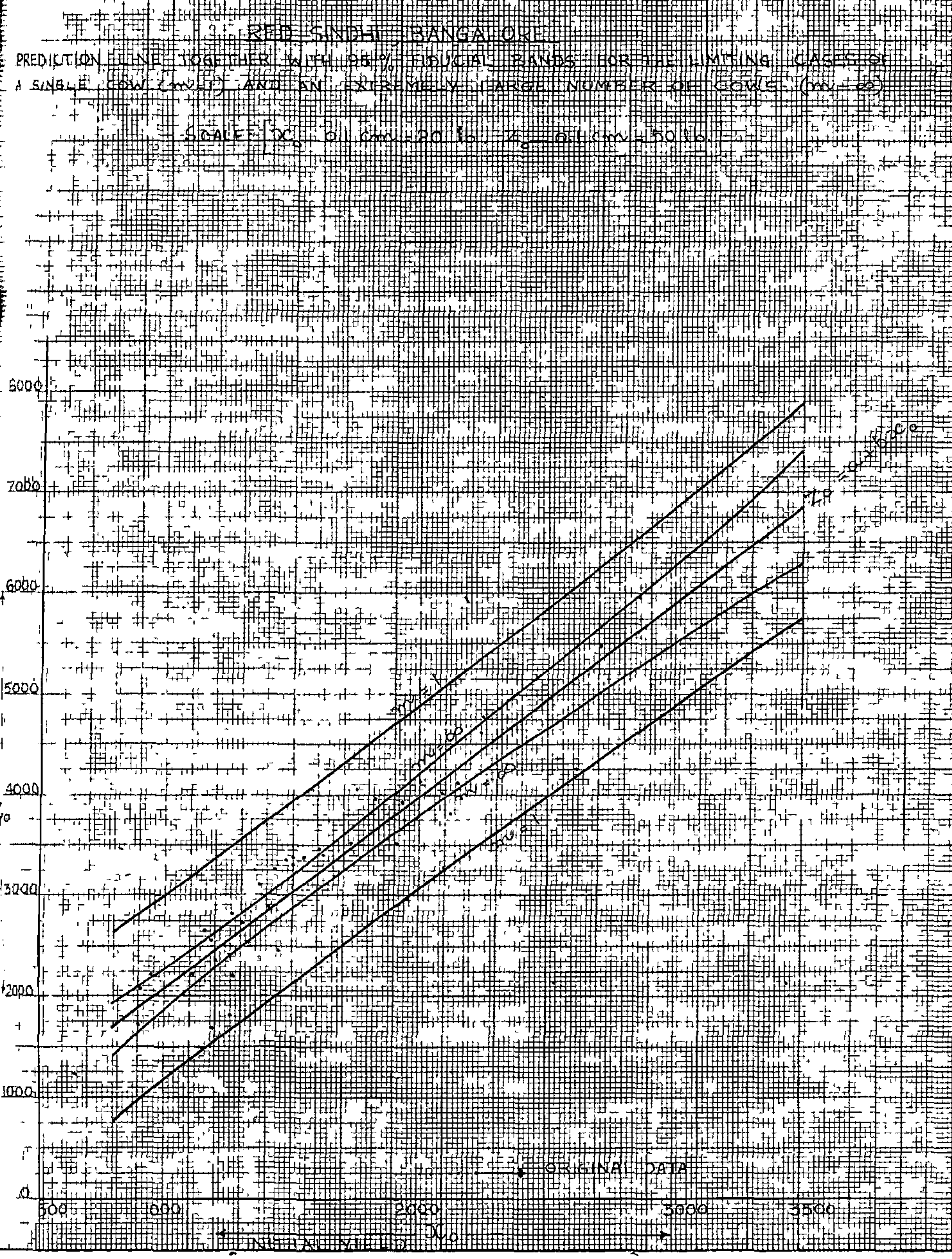

RED SINDHI BANGALORE
PREDICTION LINE TOGETHER WITH 95% FIDUCIAL BANDS FOR THE LIMITING CASES OF A SINGLE COW (m=1) AND AN EXTREMELY LARGE NUMBER OF COWS (m→∞)
SCALE
8000
7000
6000
5000
4000
3000
2000
1000
0
500
1000
2000
3000
3500
ORIGINAL DATA
INITIAL YIELD

FIGURE 2.3

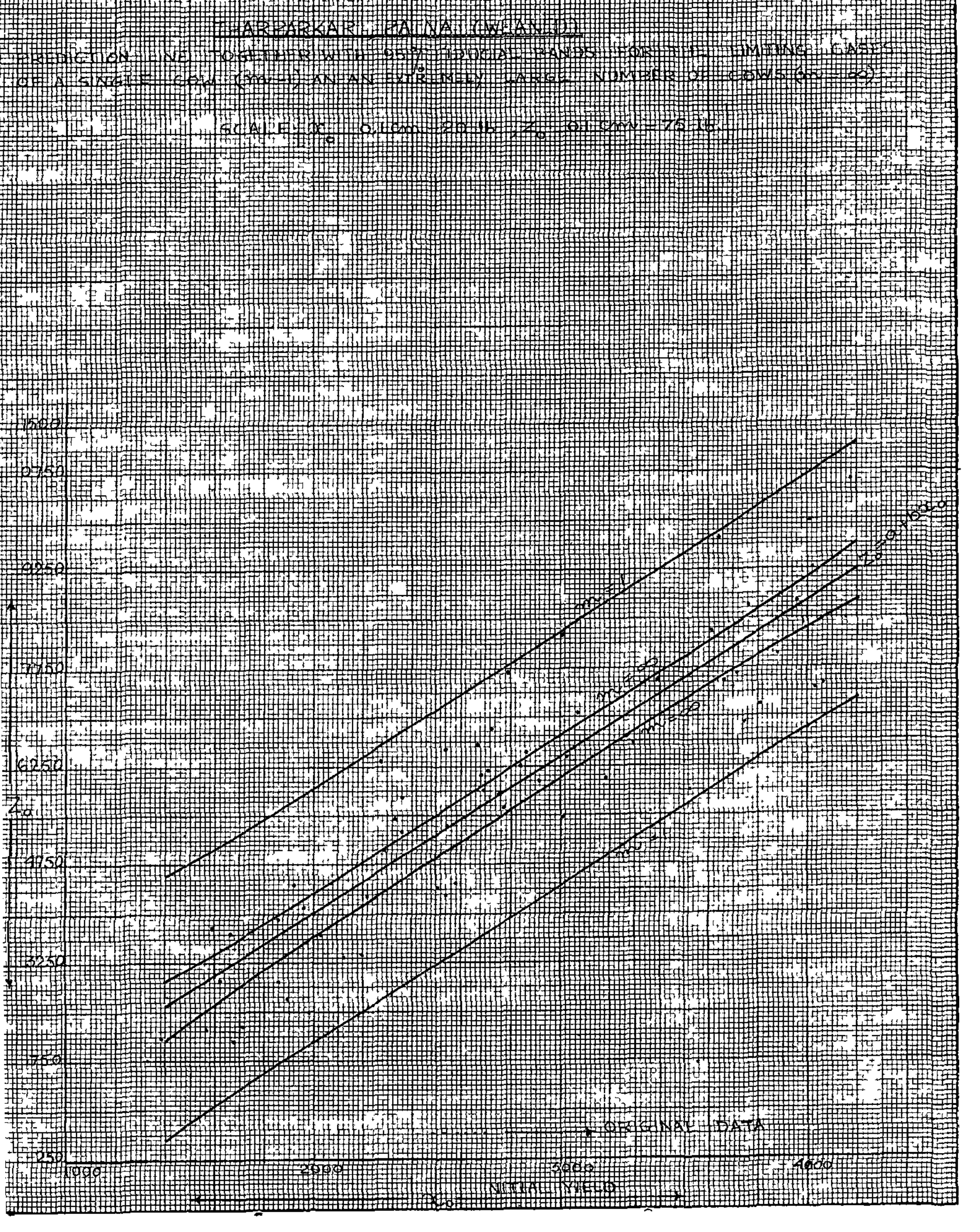

9250
7750
6250
4750
3250
250
1000
2000
3000
4000
ORIGINAL DATA
YIELD
m = 1
m = ∞

FIGURE 2.4

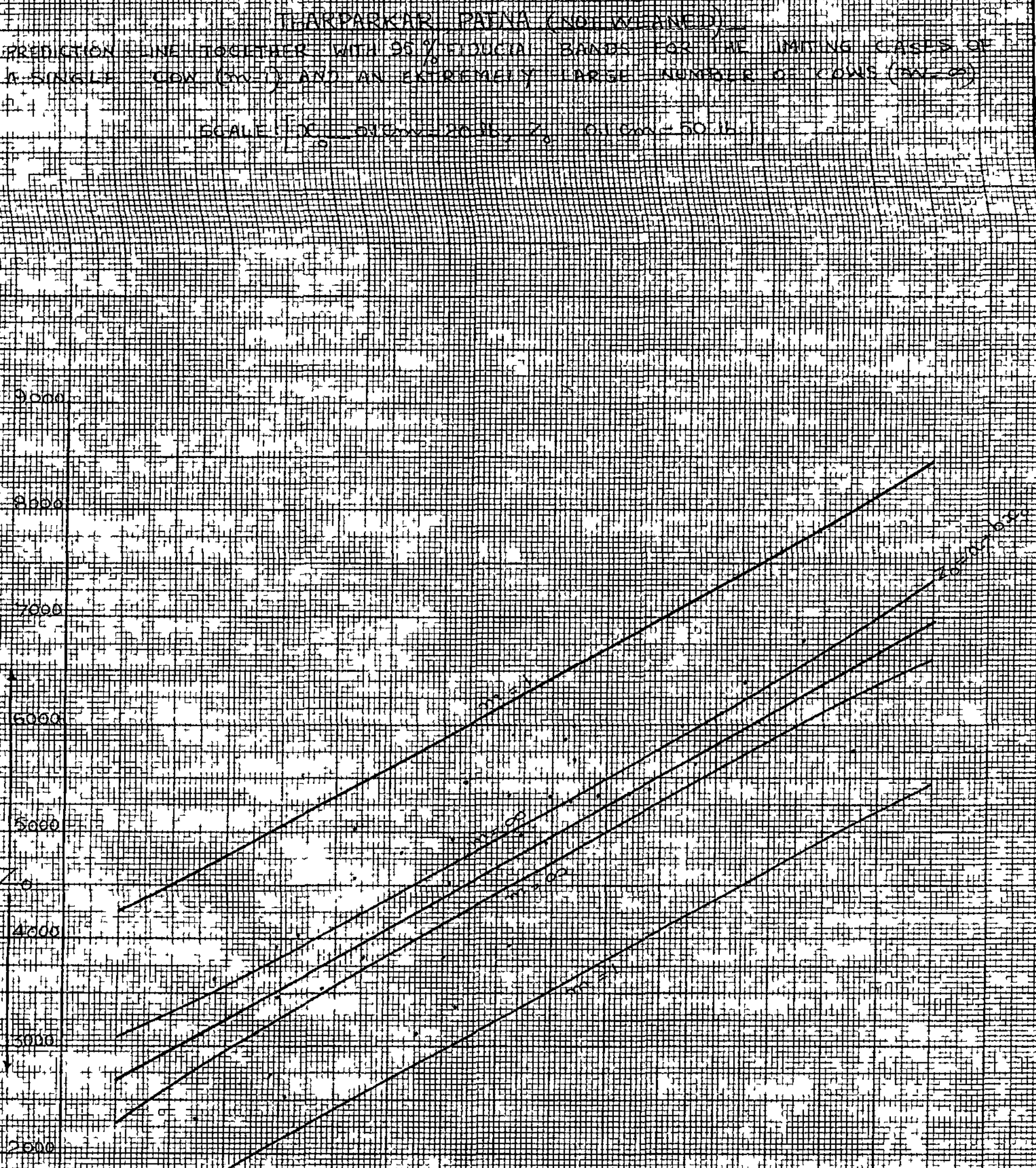

THARPARKAR PATNA (NOT WEANED)
PREDICTION LINE TOGETHER WITH 95 % FIDUCIAL BANDS FOR THE LIMITING CASES OF A SINGLE COW (m=1) AND AN EXTREMELY LARGE NUMBER OF COWS (m=∞)
SCALE: x_0 0.1 cm = 20 lb, Z_0 0.1 cm = 50 lb
9000
8000
7000
6000
5000
Z_0
4000
3000
2000
1000
1800
2000
2500
3000
3500
4000
4500
5000
ORIGINAL DATA
x_0 INITIAL YIELD

FIGURE 2.5

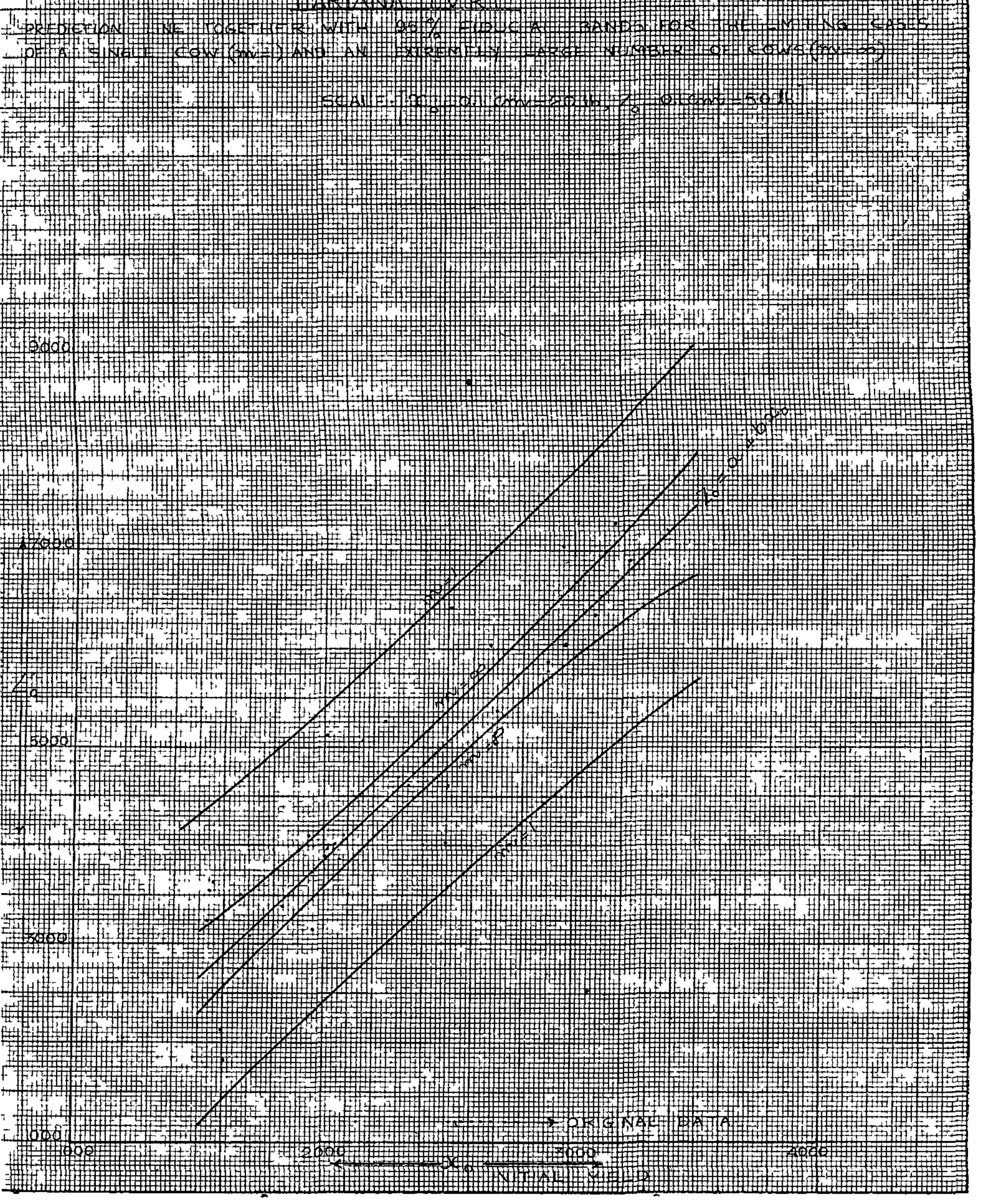


FIGURE 2.6

HARIANA, HISSAR.

PREDICTION LINE TOGETHER WITH 95% FIDUCIAL BANDS FOR THE LIMITING CASES OF A SINGLE COW (m=1) AND AN EXTREMELY LARGE NUMBER OF COWS (m=∞)

SCALE: x_0 1 cm = 10, Z_0 1 cm = 10 lb.

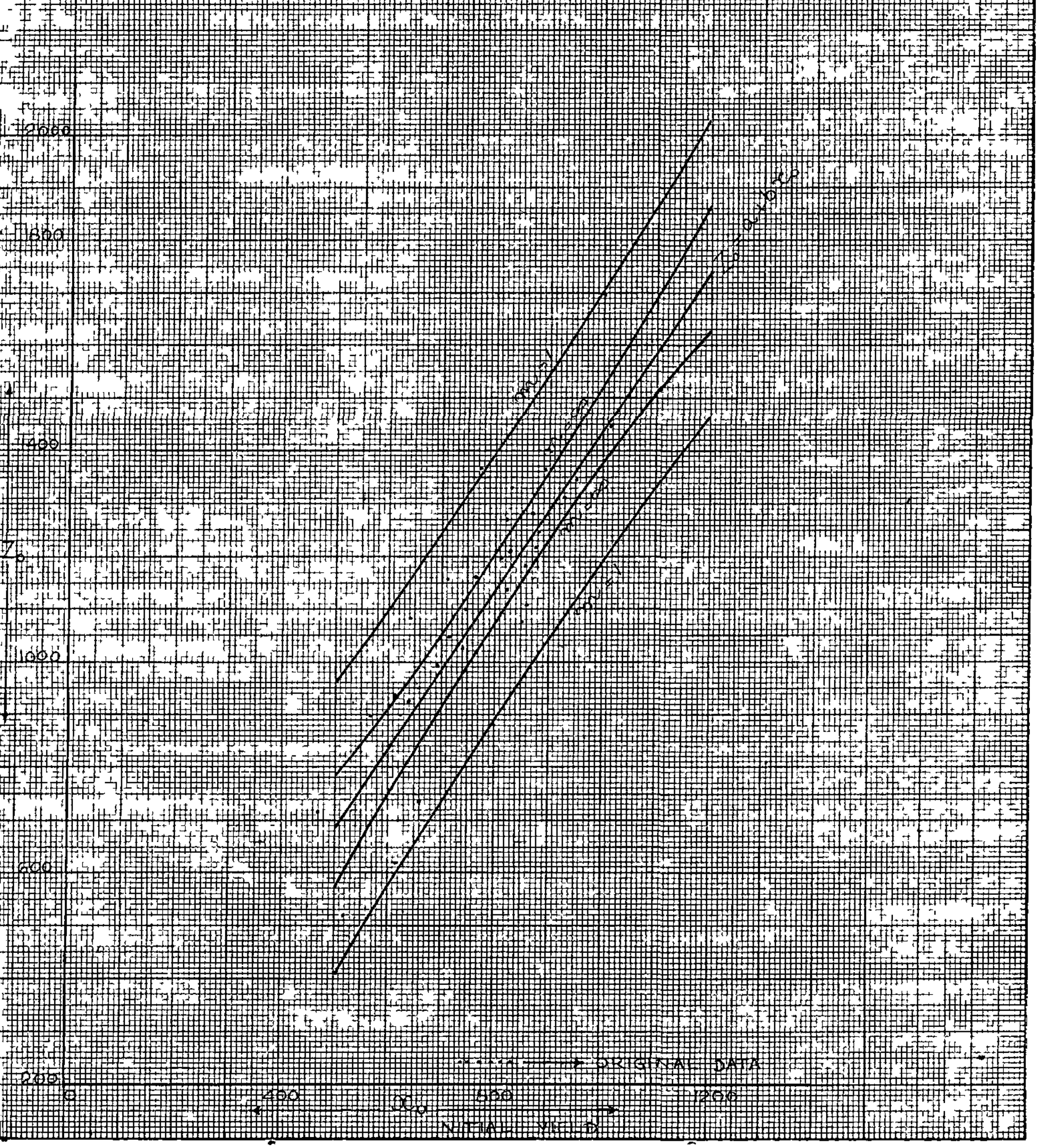
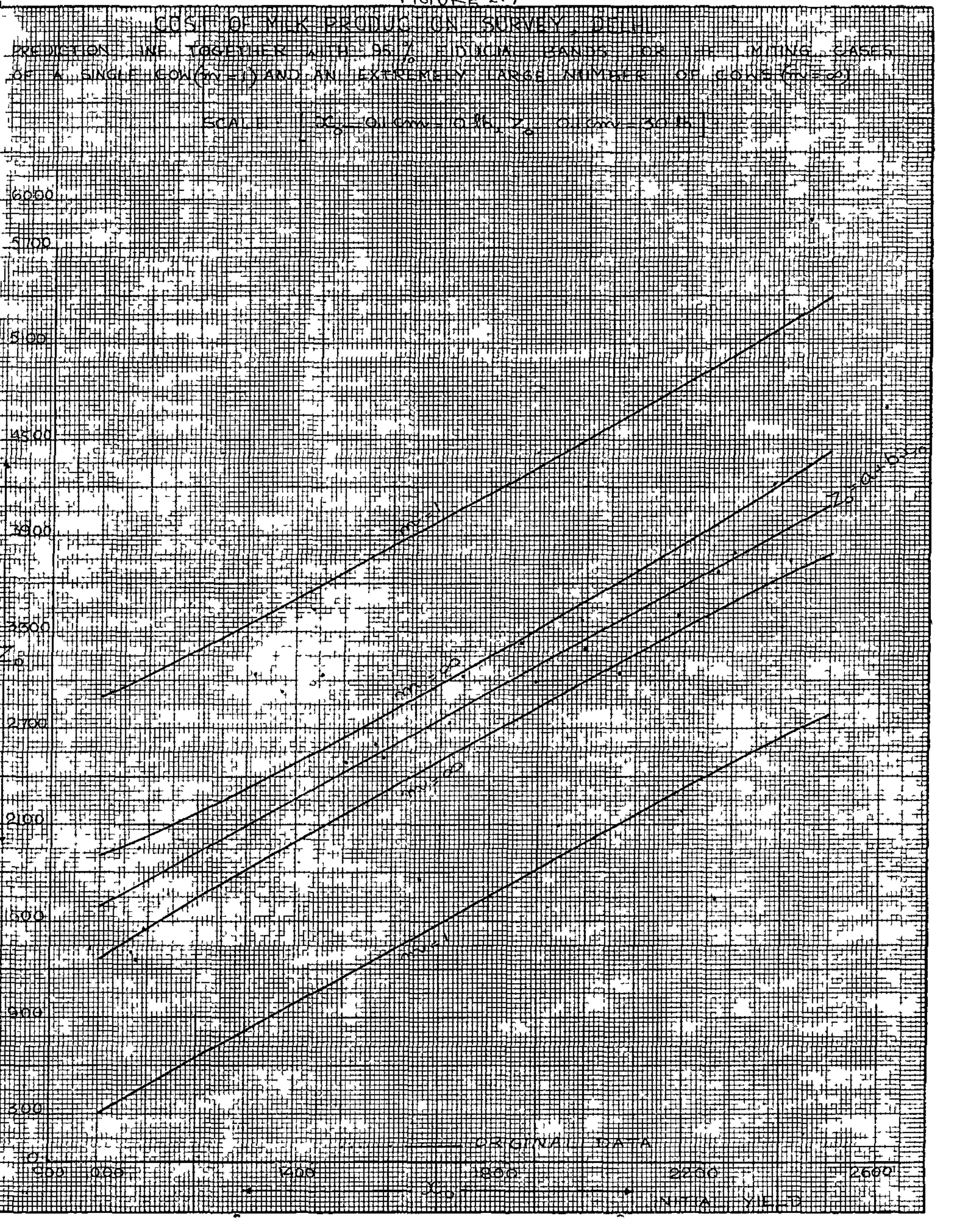


FIGURE 2.7


COST OF MILK PRODUCTION SURVEY, DELHI
SCALE : | x0 – 0.1 cm = 10 lb, Z0 0.1 cm = 30 lb |
6000
5700
5100
4500
3900
3300
2700
2100
1500
900
300
0
900
1000
1400
1800
2200
2600
Z0
X0
ORIGINAL DATA