Vol XXIII
No. 1

JANUARY-
MARCH
1968

ISSN 0019-5014

INDIAN JOURNAL OF AGRICULTURAL ECONOMICS

INDIAN SOCIETY OF
AGRICULTURAL ECONOMICS,
BOMBAY

NOTES

AN INVESTIGATION ON THE ERRORS IN THE ESTIMATES OF YARDSTICKS OF PRODUCTION FROM FERTILIZERS USE

It is proposed to increase the food production during the Fourth Five-Year Plan by 28 million tonnes over the Third Five-Year Plan. The major contribution to this additional production is to come from the use of fertilizers. In order to set up realistic targets of production from the use of inputs, such as, fertilizers, it is necessary to have a knowledge of the response per unit of the input. Such relationship between output and the input has been termed as yardstick of production for purpose of planning.

The estimates of yardsticks of production for fertilizers and manures have been worked out for several crops.[1] These yardsticks were based on the results of experiments conducted in a sample of cultivators' fields for a period of generally 3 to 4 years. Consequently, the estimates are subjected to sampling errors. Further, there are certain non-sampling errors also involved in the use of the calculated yardsticks for estimating the additional production from the use of fertilizers. The sampling errors and biases involved in the estimates of yardsticks of fertilizers are discussed in the present paper.

Definition of Yardstick

The main purpose of calculating a yardstick is to estimate the expected change in the production from a specified change in the quantity of the fertilizer input used in a State or a region. From this point of view, we may define the yardstick R as the yield which when multiplied by the units of the additional input of fertilizer gives the estimate of the additional expected production. Thus, if Q_t and Q_o are the quantities of fertilizer, say, nitrogen used at time t and at time o (base period) respectively, P the additional production, then $P = R\,(Q_t - Q_o)$.

Therefore, $R = \frac{P}{Q_t - Q_o}$ which we may designate as the true yardstick. In practice, we calculate a quantity, say, R' which approximates to the true yardstick as closely as possible.

Nature of Data Used and Method Adopted for Calculating Yardsticks from Fertilizers Use

The yardsticks given in the study referred above were calculated based on the data of extensive fertilizer experiments conducted in scattered cultivators' fields in various districts, the fields for the experiments being selected randomly after stratification of the district suitably. The trials which consisted of application of various combinations and levels of nitrogen, phosphorus and potash were superimposed on the normal practices of farmers so that the responses obtained could be considered directly applicable under farmer's conditions. Quadratic response curves were fitted to the responses from these trials in each State and the responses

1. V. G. Panse, T. P. Abraham and C. R. Leelavathi: Yardsticks of Additional Production of Certain Foodgrains, Commercial and Oilseed Crops, Institute of Agricultural Research Statistics (I.C.A.R.), New Delhi, 1964.

to 22.4 kg. per hectare of the nutrient were estimated from the average response curve. This response divided by the dosage is taken as the yardstick.

The sources of error that are encountered in estimating yardsticks in the above manner are :

(1) Sampling error which arises due to estimating the parameters only from a sample of the fields and not from all the fields. Further, the variation due to seasons also adds to this error.

(2) Specification error which is due to the failure of the model to reflect accurately the conditions actually obtained in the farm and also errors in the statistical specification of the model, such as, the response curves.

(3) Aggregation error defined as the difference between the aggregate additional production as developed from the summation of production from individual fields and that developed from the average response curve.

The data utilized in the calculation of yardsticks are those obtained from cultivators' fields and, as such, can be considered to reflect properly the conditions obtaining in their fields. A quadratic fertilizer response curve has been assumed in the calculation. The work done with extensive fertilizer trial data in the country has shown that this function gives an adequate representation of the responses.[2] Therefore, in the subsequent discussion only sampling errors and aggregation errors have been mainly considered. It may, however, be observed that in the calculation of yardstick a standard dose has been taken for calculating the response. This standard dose was 22.4 kg. per hectare (20 lbs. per acre) for cereals. The error introduced due to this standardization has, however, been taken into account in the discussion.

Sampling Error of the Estimated Yardstick

The sampling error of the estimates is composed of the two components: (*a*) error arising out of the sampling of fields (σ_e^2) and (*b*) error due to year to year differences in the responses (σ_b^2). These two components of variation are estimated from the analysis of variance of the responses in each year in each district and are given in Tables I and II for nitrogen and phosphorus on rice and irrigated wheat. The year component of variation for nitrogen yardstick on rice is quite large, particularly, in Kerala, Andhra Pradesh and Assam. On the average, the year to year variation is nearly 60 per cent of the variation within the district in a given year.

The average year component of variation is larger than the variation between fields within district for irrigated wheat. Generally, for the latter crop, the variance components were higher than for rice. In the case of phosphate, the year component was relatively smaller compared to within district variation. This shows that while response to phosphate is more variable from field to field, it is more stable from year to year compared to nitrogen, indicating that seasonal factors exert less

2. T. P. Abraham and V. Y. Rao, An Investigation of Functional Models for Fertilizer Response Surfaces, *Journal of the Indian Society of Agricultural Statistics*, Vol. 18, No. 1, 1966.

TABLE I—NITROGEN YARDSTICK
COMPONENTS OF VARIANCE, YARDSTICK, STANDARD ERROR AND CONFIDENCE INTERVAL

Crop	State	Yard-stick in kg. per kg. of nitrogen	Standard error in kg./kg.*	95 per cent confidence interval in kg./kg.	σ_e^2 in kg. of response per kg. of nutrient	σ_b^2 in kg. of response per kg. of nutrient
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rice	Andhra Pradesh	10.75	1.70	7.41 to 14.11	3.4275	2.2335
	Assam	10.71	1.91	6.96 to 14.46	2.2496	2.3230
	Bihar	11.61	0.60	10.40 to 12.81	1.4330	0.2681
	Kerala	12.10	3.06	6.12 to 18.08	0.9264	7.4163
	Madras	10.40	0.53	9.37 to 11.43	5.9668	0.0000
	Madhya Pradesh	13.08	1.57	10.00 to 16.16	4.7410	1.3797
	Mysore	11.29	0.86	9.60 to 12.99	4.2087	0.4548
	Punjab	12.14	1.24	9.69 to 14.60	3.3026	1.1049
	Uttar Pradesh	10.22	0.40	9.42 to 11.03	0.8618	0.1155
	West Bengal	8.57	1.02	6.56 to 10.58	1.3843	0.7428
Irrigated Wheat	Bihar	12.77	1.70	9.46 to 16.07	3.0293	2.2281
	Punjab	16.16	2.97	10.36 to 21.96	5.5026	6.9234
	Rajasthan	9.78	1.91	6.03 to 13.53	4.2928	2.7134
	Uttar Pradesh	15.89	0.95	14.02 to 17.77	1.2276	0.7015

* Standard error underlying confidence interval.

TABLE II—PHOSPHATE YARDSTICK
COMPONENTS OF VARIANCE, YARDSTICK, STANDARD ERROR AND CONFIDENCE INTERVAL

Crop	State	Yard-stick in kg. per kg. of nitrogen	Standard error in kg./kg.*	95 per cent confidence interval in kg./kg.	σ_e^2 in kg. of response per kg. of nutrient	σ_b^2 in kg. of response per kg. of nutrient
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rice	Andhra Pradesh	6.70	1.02	4.68 to 8.70	2.7863	0.7663
	Assam	10.40	0.89	8.66 to 12.14	4.8429	0.3885
	Bihar	8.44	0.82	6.83 to 10.04	1.6455	0.5025
	Kerala	9.37	0.20	8.97 to 9.78	0.4747	0.0000
	Madras	7.45	1.97	3.57 to 11.34	6.6479	2.6815
	Madhya Pradesh	8.39	2.61	3.30 to 13.48	6.0120	4.1939
	Mysore	8.26	1.40	5.49 to 11.03	4.5488	1.4684
	Punjab	6.07	1.48	3.17 to 8.97	2.7376	1.6489
	Uttar Pradesh	5.80	0.66	4.51 to 7.10	0.8086	0.3123
	West Bengal	3.84	1.01	1.87 to 5.80	1.2523	0.7294
Irrigated Wheat	Bihar	9.37	0.28	8.84 to 9.91	2.7941	0.0000
	Punjab	8.26	0.44	7.41 to 9.11	7.6012	0.0000
	Rajasthan	7.50	1.45	4.64 to 10.36	5.0581	1.4259
	Uttar Pradesh	9.55	0.71	8.17 to 10.94	1.4768	0.3937

* Standard error underlying confidence interval.

disturbing influence on phosphatic response. The standard error of the estimated yardstick varies from 2.18 per cent in Uttar Pradesh to 10.83 per cent in Assam with nitrogen on rice and from 2.81 per cent in Uttar Pradesh to 10.04 per cent in Rajasthan on irrigated wheat. The variation with phosphorus ranged from 1.90 per cent in Kerala to 19.15 per cent in Madhya Pradesh on rice and from 2.86 per cent in Bihar to 10.71 per cent in Rajasthan on irrigated wheat.

Confidence Interval of the Estimated Yardsticks

The yardsticks are used for predicting the yield response in a future year. It will, therefore, be of interest to calculate the confidence interval in using yardstick for this purpose.

Using the estimates of the components of variance σ_b^2 and σ_e^2 the 95 per cent confidence interval for the true yardstick of lth year can be approximately put as

$$\bar{x}_{mn} - 1.96\sqrt{\hat{\sigma}_b^2\left(1+\frac{1}{n}\right)+\frac{\hat{\sigma}_e^2}{mn}} \leqslant \mu.l \leqslant \bar{x}_{mn} + 1.96\sqrt{\hat{\sigma}_b^2\left(1+{}^n\right)+\frac{\check{\sigma}_e^2}{mn}}$$

where 1.96 is the 95 per cent point of a standardized normal deviate, $\bar{x}_{mn}$ the mean yardstick based on m districts and n years. The calculated confidence intervals are also given in Tables I and II.

The 95 per cent confidence intervals for nitrogen yardstick on rice are comparatively large for all the States except Bihar, Madras, Mysore and Uttar Pradesh. The same is true for irrigated wheat also except in Uttar Pradesh. The confidence interval of yardstick of phosphorus except for Kerala on rice and Bihar for wheat are also very large.

From the magnitude of the sampling errors and the corresponding confidence interval of the estimated yardsticks it appears that in using these yardsticks for estimating the production for any particular year, there is a relatively large component of error involved. However, the margin of error is considerably reduced when it is used to estimate the production for a period of years.

Bias due to Specification and Aggregation Errors

The bias introduced in the yardstick due to the combined error of aggregation and differences in the doses actually used by farmers from the standard dose used in the estimation of yardstick can be shown to be approximately.

$$\gamma\left[s - \frac{\sigma_t^2 - \sigma_o^2}{\mu_t - \mu_o} - (\mu_t + \mu_o)\right] + \frac{1}{\mu_t - \mu_o}\left[\text{cov}\,(b, \triangle x) - \text{cov}\,(c, x_t^2) + \text{cov}\,(c, x_o^2)\right] \quad \ldots \quad (1)$$

where x_t and x_o are the rates of application in the tth year and base year respectively, b the linear coefficient of the fertilizer yield regression equation, c the corresponding quadratic coefficient, $\triangle x$ the difference in $x_t - x_o$, μ_t and μ_o the expected values of x_t and x_o and Υ the expected value of c. We have assumed that the rate of application is not correlated with the size of the field. If we further assume that the rates of application are uncorrelated with the magnitudes of the b and c coefficients, the bias will become

$$\Upsilon\left[s - \frac{\sigma_t^2 - \sigma_o^2}{\mu_t - \mu_o} - (\mu_t + \mu_o)\right] \quad \ldots \quad \ldots \quad \ldots \quad (2)$$

Therefore the bias will vanish if s, the standard level is taken as

$$\frac{\sigma_t^2 - \sigma_o^2}{\mu_t - \mu_o} + (\mu_t + \mu_o) \quad \ldots \quad \ldots \quad \ldots \quad (3)$$

Thus even if we assume no correlation between the rates of application and the magnitudes of the response curves, the calculation of the bias will require a knowledge of the mean rates of application in the base year and in the current year as well as the corresponding variances of the rates of application in addition to the expected value of the quadratic component of the response function. The mean rates of application can be easily calculated given the total quantities of the fertilizers to be applied in the current year and what actually was applied in the base year, for, these quantities divided by the corresponding area under the crop give the mean rates of application. Here again, while information on the fertilizer quantities will be available for all crops together, it will be difficult to get information on the fertilizer consumption for each crop separately. Even more difficult is to get an estimate of the variance of rates of application which can be obtained only by undertaking sample surveys to collect information on the rates of application. If all this information was available, we could practically remove the bias by suitable choice of the standardization level for calculation of yardstick. The Institute of Agricultural Research Statistics has carried out some sample surveys in selected districts to study the fertilizer practices of the farmers. In a few districts, repeat surveys were also carried out after a period of 5 to 6 years. The data provide information on the rates of application of fertilizers on various crops which could be utilized to examine the nature of bias involved in using the yardstick.

The details of the surveys are given in the report on Farmer's Fertilizer Practices.[3] Information was collected on the rates of application of fertilizer from each holding. The mean rates of application together with the mean and standard deviation are given in Table III for nitrogen application in Coimbatore and Barabanki districts where repeat surveys were carried out. Coimbatore district represents an area where the rates of application of fertilizer have been high while Barabanki, located in Uttar Pradesh, has comparatively lower rate of fertilizer. It will be of interest to work out the bias involved in these two extreme cases. The biases involved in the yardstick, taking first year of the survey as the base year and the repeat survey as the last year, are given in Table III. The bias has been calculated using formula (2).

3. Indian Council of Agricultural Research: Farmer's Fertilizer Practices: Coimbatore (Madras) 1954-55, and Barabanki (Uttar Pradesh) 1956-57.

TABLE III—ESTIMATED BIAS OF YARDSTICK FOR NITROGEN

District	Crop	Year	Mean rates of application kg. N/ha.	Standard deviation of rates of application kg.N/ha.	Bias
Coimbatore	Rice	1954-55 1961-62	24.0 35.7	23.9 29.9	+ 7.8
Barabanki	Rice	1956-57 1962-63	6.8 9.1	8.9 10.2	+ 0.5
	Wheat	1956-57 1962-63	5.4 6.6	12.1 14.3	+ 4.6

Standardisation level 22.4 kg. N/ha.
$\gamma = 0.12$ per kg. N and kg. response.

The expectation of the quadratic coefficient of the response function has been taken as 0.12, the units being kg. of output and kg. of nutrient. This value has been taken since extensive fertilizer trials in cultivators' fields showed that by and large the average c value (quadratic coefficient) of the response function is of this order. It will be noted that in the case of Coimbatore, where the rates of application even in the base year have been higher than the standard level taken for calculation of yardsticks, the bias in the use of yardstick is very high being an overestimate by about 8. At Barabanki, similar calculation shows for rice the bias of the order of only 0.5 while for wheat in the same district, the bias is nearly 5 units which should be considered quite large since the yardsticks are generally of the order of 10 to 14 units. Thus we find that in Barabanki even though the rates of application are not much different for rice and wheat, the bias due to aggregation is much larger for wheat due to larger variability in its rate of application.

The analysis carried out above indicate that apart from sampling error, the biases introduced by aggregation and standardization of level are quite considerable in some cases. With low rates of application, as in the case of Barabanki, the error introduced is not likely to be large, but with rates of application as large as in Coimbatore, the situation is different. It will be desirable to use different standardization levels in the calculation of yardsticks for different areas depending on the pattern of fertilizers use. In order to use such differential yardsticks for estimation of additional production it is also necessary to obtain information on the targets of distribution in these areas. With rapid increase in the use of fertilizers if we have to make realistic estimates of the additional production from fertilizer use, these changes regarding the standardization level and the areawise targetting will be essential.

Summary

Yardsticks of production from fertilizer use have been given by Panse, Abraham and Leelavathi.[4] In the present paper, the sources of error that are encountered in the estimation of these yardsticks have been examined. These

4. *loc. cit.*

errors are the sampling error, the specification error and the aggregation error. The standard error of the estimated yardsticks varied from about 2 to 10 per cent generally.

The formula for bias involved has been worked out under certain assumptions. It was found that the bias due to aggregation and standardization can be considerably reduced if the standardization level is chosen as somewhat higher than the sum of the average rates of application in the base year and the current year. The calculations are illustrated with the data obtained in fertilizer survey carried out in a few selected districts. It has been brought out that it is essential to use different standardization levels in different areas if substantial errors in the use of the yardsticks have to be avoided.

T. P. ABRAHAM
AND
C. R. LEELAVATHI†

MEASUREMENT OF SIZE OF FARM AND EFFICIENCY—SOME ALTERNATIVE APPROACHES*

Objective

The article attempts to examine the definition of the 'size of farm' and to show that measures of size and efficiency other than those currently used for measuring size (namely, geographical area) and efficiency (namely, the yield per acre) lead to different conclusions. An analysis of this problem has also been made by Raj Krishna.[1]

SIZE OF FARM

In practically all the available studies relating to agriculture whether by census of land holdings, cost data or input/output analysis, the size of the farm is invariably measured by its geographical area. The size of a farm can, in fact, be measured either in terms of physical output, gross receipts (total and per acre), total inputs, etc. It is of importance to choose the concept of size correctly, because the size distributions of farms according to different measures of size are not similar.

The classifications of farms according to the land and labour input would be useful provided that classifications by value added and by each of the other major input groups such as, say, seeds, manures, fertilizers, etc., are also given. But such data are rarely available together. However, classification according to inputs only, ignoring the aspect of output, is not wholly satisfactory. This flows from the fact that the scale distribution proper and the effect of relative efficiency of different farms are mixed up. This prevents identification of size and efficiency magnitudes and analysis of the relation between them.

† Assistant Statistical Adviser and Statistician, respectively, Institute of Agricultural Research Statistics (Indian Council of Agricultural Research), New Delhi-12.

* The views expressed in this paper are the personal views of the author and not necessarily of the organization he has the privilege to serve.

1. Raj Krishna, "The Optimum Firm and the Optimum Farm," *The Economic Weekly*, Vol. XIV, Nos. 40 and 41, October 6 and 13, 1962.