THE ACCURACY OF YIELDS OF GRASSLAND AND OATS EVALUATED BY EYE ESTIMATES [1])

TH. J. FERRARI

Agricultural Experiment Station and Institute for Soil Research T.N.O., Groningen

SUMMARY

In agriculture estimates are often used. In this respect an estimate means the result of determining a quantity in which method either a scale is used imprinted on the mind (mental estimate) or the quantity is derived indirectly by means of a number of secondary characteristics (correlative estimate). The use of estimates has its advantages especially in economizing time and money; sometimes it is the only method of determining the characteristic (e.g. the quality).

The usefulness of the estimates is determined by the degree in which they represent the real value. A method is given to determine the standard error of eye estimates of the yields of grassland and oats. The estimates given in relative values running from 1 to 100 must be converted into absolute values. The calibration and the investigation into the accuracy of the eye estimates belong to the field of line fitting in which the weighed values and the estimates both are subject to error. The total errors are splitted up in accidental errors and in systematic errors on account of the subject and on account of the object. It is possible to diminish only the first two errors by replicated observations.

1 INTRODUCTION

In the Netherlands especially VISSER (1948, 1951) has pointed out the significance estimates can have in solving questions of a scientific or practical nature in agriculture. In this regard an estimate is understood to be the method of determining a quantity, whereby either – this will be mostly the case – a scale is used which in general is more or less imprinted on the mind (mental estimate) or the quantity is indirectly derived via a number of secondary characteristics, possibly to be determined with the aid of mental estimates (correlative estimate).

Provided a number of conditions are complied with, advantages of various kinds, not the least of which is a saving of time and money, are connected with estimates of different characteristics. A determination of the yield of cereals by way of an eye estimation can be effected about 25–30 times quicker than a determination of the yield by means of weighing. The significance of the former method is further enhanced in determinations of taste and quality, for which no objective method of determination has been evolved. On the other hand some objections are also attached to estimates, resulting from the special character of the estimates. These objections, however, do not prevent estimates being employed in many cases. POSTHUMUS (1943), MURRAY (1952), HOCKENSMITH (1948), YATES (1949), EZEKIEL (1950) and ROTHKEGEL (1952), mention the employment of estimates as a method of determination.

It is clear that eventually the significance of the estimate has to be judged from the extent to which it gives a picture of the real value. This is after all a requirement which applies to every determination.

In general not many objections will be raised when the supposition is made

[1]) Received for publication March 18, 1953.

that it is possible to realize more or less clearly the differences (e.g. in the yield) by estimates; a considerable part of everyday life is focussed on this possibility. The degree of accuracy determines however its practicability. In view of the subjective element inherent in the mental estimate a regularly repeated check of the comparability of different sequences of estimates is essential.

The problem in correlative estimates is to indicate one or more characteristics in such a way that the quantity required can be deduced from them in a certain way with a degree of accuracy which meets the requirements. Also in the case of a correlative estimate a check of its accuracy will have to be decisive as regards its practicability. In many cases checks will have to be repeated at regular intervals, as it is possible that the correlations employed change under altered conditions (e.g. the year).

It is surprising that most of the previously mentioned authors pay little attention to the accuracy of mental estimates. Only Posthumus (1943) gives a circumstantial study of the accuracy employed in organoleptic butter appreciation; he employed a method whereby the subjective "ideal" estimate (to be approached by taking a large number of persons who make an estimate) is taken as a standard. We present a method to determine the accuracy of the eye estimate of the yields of grassland and oats by comparison with weighed yields.

2 Some reflections on the investigations into the accuracy of mental estimates

The best way to arrive at a judgment of the accuracy of mental estimates is by comparing the estimated value with the real value. The difficulty in this respect is however that the real value is usually unknown because weighing and measuring are subject to errors in observation.

The accuracy can be checked from the *error* of the determination. The idea error can in this respect be said to be the deviation to be expected between the estimated and the real value of the characteristic to be determined. The solution of these problems of determination of accuracy lies in the field of mathematical statistics.

The divergence found in the previously mentioned accuracy investigations is therefore not only caused by the error of estimate but also by the error of weighing and measuring. We can therefore differentiate in this respect:

a. *The accidental error of weighing and measuring.* The extent of this error in every individual determination in a sequence of repeats is unknown; the influence on the ultimate result, for instance on the average, can be ascertained.

b. *The accidental error of estimate.* The same applies to this concerning the accidental error of weighing and measuring as is stated under a.

c. *The systematic error of estimate.* The extent of this error in every individual estimate in a sequence of repeats is the same; the influence on the ultimate result, for instance on the average, can be determined. The origin of this error depends on the kind of estimation. In certain cases it is desirable to differentiate the systematic error of estimate in a systematic error *on account* *of the object* and in the systematic error *on account of the subject.* The former error is the same in every individual estimate in a sequence of repeats, effected by different valuers estimating independently of each other; it is due to the fact that the object has properties which make all these valuers estimate this object, for instance the yield in a certain field, equally too high or too low. The latter error is the same in every individual estimate in a sequence of repeats, now effected by the same valuer.

Experience shows that the systematic error of estimate has in general the greater significance; its extent eventually determines the possibility of estimating a certain characteristic. Every method to determine the total error of estimate will have to be suitable to approximate the systematic error.

In many estimates (for instance of the yield) a relative scale is used. It is desirable also to consider the estimates, where an absolute scale is used, as relative estimates; for further investigation will have to show to what extent the absolute scale has been correct. The check of the accuracy of the estimates therefore belongs to the field of line fitting, whereby for the determination of the constants of the line of adjustment the errors of the estimation as well as those of the weighing and measuring have to be taken into account.

It is clear that one can stipulate that the scales of the different sequences of estimates are equal and that a scale remains the same along the whole course; the execution and the adjustment are also considerably simplified when this requirement has been complied with. Mathematically these requirements mean that calibration lines are rectilinear and that the different lines have the same tangent. It is experimentally proved that these requirements cannot always be complied with, though it is feasible that with a proper control and instruction much can be achieved. It appears that in favourable cases rectilinearity is achieved but that the scales of the valuers are different, i.e. that the gradients of the calibration lines are different in size. It is also often experienced that the rectilinearity disappears in certain stages owing to high values being estimated too low, low ones too high. The consequence is that not only the determination of the line of adjustment and of the error of estimate are considerably impeded, but also that the error of estimate is increased. For, the error of estimate is converted by means of the calibration line into weighing or measuring units; a deviation in the scale as previously mentioned and manifesting itself in a curvilinear relation means essentially that an estimation unit conforms to a larger number of weighing units.

Summary: it is appreciated (in order of decreasing importance) that the total error of estimate be as small as possible and that the graph be rectilinear, preferably with a constant gradient and that it also runs through the origin of the system of co-ordinates.

3 The determination of the error of estimate

The problem can therefore be put thus: that we surmise that between the measured or weighed value x and the estimated value y there is a linear relation $y = ax + b$; x and y refer to the same quantity, they are only expressed in different scales, which shows that the line is the calibration line. The observation of the points on this line co-incides with errors in observation of the estimate as well as of the measured or weighed value. The question is to calculate the constants a and b of the line and the extent of the error in observation of one of the methods of determination or both with the aid of the observations (x_i , y_i), whereby i = 1, 2,. . . .n.

The problem of the line fitting when both variables contain errors in observation has been dealt with by various investigators. For a survey we refer to HEMELRIJK (1952) who discusses the most suitable older methods side by side with the newer ones. It is a general principle for a choice from the different methods that a better result will be obtained as more assumptions are used. The question is then however whether the assumptions comply readily with the requirements. The material at our disposal the error of which in observation of the yield measuring was known, enabled us to check the results obtained with the method of WALD (1940) and DRION (1948) ; with these methods it is possible not only to calculate a line of adjustment but also the errors in observation of the two variables. The agreement between the well-known errors in observation of the weighing and those calculated in this way was very small though we tried to make the observation material comply as much as possible with the requirements.

We have used the method we give hereafter which makes it possible to calculate the error of estimate with the aid of the error in observation of the measured yield, known from elsewhere. Data will in general comply with requirements ; errors must be independent of the real value, observations further have to comply with conditions made for the calculation of the correlation co-efficients.

The consequence of the errors in observation is that the square of the standard deviation s of a frequency distribution is increased according to :

$$s'^2 = s^2 + s^2_w , \qquad 1$$

in which s_w is the standard error in observation.

The relation taken as s'^2/s^2_w = p, makes the formula 1 into :

$$s^2 = s'^2 . \frac{p-1}{p} . \qquad 2$$

It is possible to prove (SITTIG, 1949) that the correlation co-efficient r' is increased by a correction of the influence of the errors in observation of the two variables as follows :

$$r = r' \sqrt{\frac{p_x \cdot p_y}{(p_x - 1)\ (p_y - 1)}} \qquad 3$$

in which p_x and p_y are the ratios of the squares of the standard deviations s' of the frequency distributions to the squares of the errors in observation s_w of x and y respectively.

We have already discussed that the deviations of the dots from the calibration line are *exclusively* caused by the occurrence of errors in observation of the two variables. Hence a correction for the influence of the two errors in observation *must* increase the correlation co-efficient to the maximum value 1, so that the formula 3 becomes :

$$1 = r' \sqrt{\frac{p_x \cdot p_y}{(p_x - 1)\ (p_y - 1)}} . \qquad 4$$

We assumed that the error in observation of one of the variables is known.

The only unknown item in formula 4 is the square of the error in observation of the other variable; the required error in observation can therefore be calculated from formula 4.

The calibration line is then determined graphically, whereby the formula for the *direction of averaging* is used:

$$\text{tg}\alpha \,.\, \text{tg}\varphi = - \frac{\text{square error in observation of } y}{\text{square error in observation of } x} \qquad 5$$

For an explanation of the tangents applied here we refer to DEMING (1946) and to figure 1.

5 THE ACCURACY OF SOME ESTIMATES OF YIELD

In the research after the soil fertility done in 1951 and 1952 by the Agricultural Experiment Station in the Guelder Valley, in which the Central Institute of Agricultural Research, the Research Department of the Government Service of Drainage, Land-improvement and Re-allocation, and the Soil Survey Institute, have shared, a number of data have been collected which makes it possible to calculate the errors of estimates of the yields of grassland and oats.

Some co-workers [1]) of the Research Department of the Government Service of Drainage, Land-improvement and Re-allocation have estimated on 50 fields of 1 are large the annual yield in kgs of dry matter per ha grassland; the scale runs from 1 to 100. In 1951 as well as in 1952 three estimates were made in order to take into account the influence of weather conditions during the season of growth, viz.: at the beginning of April, at the end of July/beginning of August, and in October. We do not go into the technique of these estimates. In order to prevent any misunderstanding we only remark that they are not estimates of the quantity of grass on the stalk; the botanical composition of the sod and the appearance of the grassland are for instance factors considered in the estimates.

The determination of the yield was effected by placing the grass under two cages, each having a surface of 5.04 m^2, and by cutting and weighing it (SCHEYGROND and SONNEVELD, 1951). The annual yield, converted into kgs of dry matter per ha, then consists of the sum of the yield of six cuts. The average error in the determination of the yield by weighing is calculated from the differences in the annual yield between the two cages according to

$s = \sqrt{\frac{\Sigma d^2}{2n}}$ where d is the difference in yield and n the number of fields.

That combination from the three estimates which gives the highest correlation co-efficient with the weighed yield is eventually employed as the best estimate. It then appears that in 1951 as well as in 1952 the average of the first and second estimate is the nearest approximation of the actual yield. Figure 1 gives the connection between the weighed yield and the average estimate in 1952.

---

[1]) The estimates were made by Messrs T. DOUMA, A. W. VAN DER POEL, A. B. HUIS IN 'T VELD and H. ZUURVEEN.

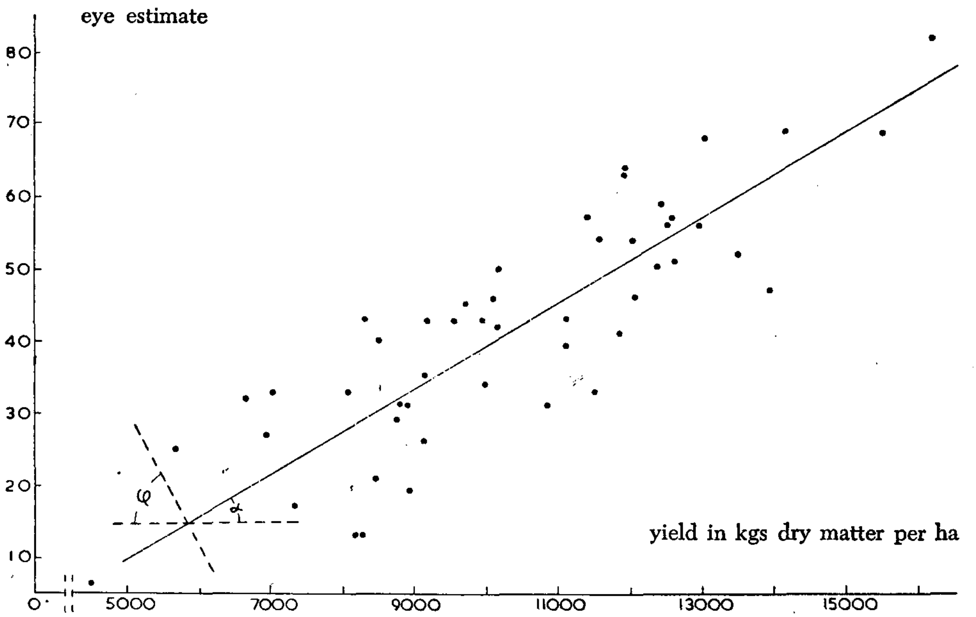

FIG. 1. RELATION BETWEEN THE EYE ESTIMATES OF THE YIELDS AND THE ACTUAL YIELDS OF GRASSLAND.

In table 1 the most important data and results of the calculations are given. The column with the data on the accidental error of estimate requires some elucidation. The fact is that we do not have any repeats to calculate the accidental error of estimate. This error is therefore approximated by taking the smallest error, calculated from the differences between two successive estimates; the correct accidental error is of course smaller, so that the systematic errors mentioned are actually larger.

Table 1 shows that the two years give corresponding results. The errors of determination of the yield by weighing as well as of the estimates may be larger in 1952 than in 1951, but, expressed in percentages of the variation amplitude, they are equal. The last column further shows the most important datum to be that the determination of the yield by estimation is only 1.7 times less accurate than the determination of the yield by weighing. This is a very striking result in consideration of the many factors which influence the yield of grassland.

Of a number of fields with oats, the yield of which was estimated, it was also determined by weighing. This number was in 1951 and 1952, 34 and 216 respectively. These 216 estimates are not independent ones; they refer to 36 experimental fields, each consisting of 6 plots.

The plots, measuring about half an are, were in 1951 a section of the practice field; in 1952 the plots were laid out and looked after by the personnel of the Agricultural Experiment Station. In contradistinction to 1951 only one oat-species, viz. *"Marne"*, was used in 1952. This uniformity has undoubtedly favourably influenced the accuracy of the estimates. The estimates were chiefly made by the co-workers of the Research Department of the Government

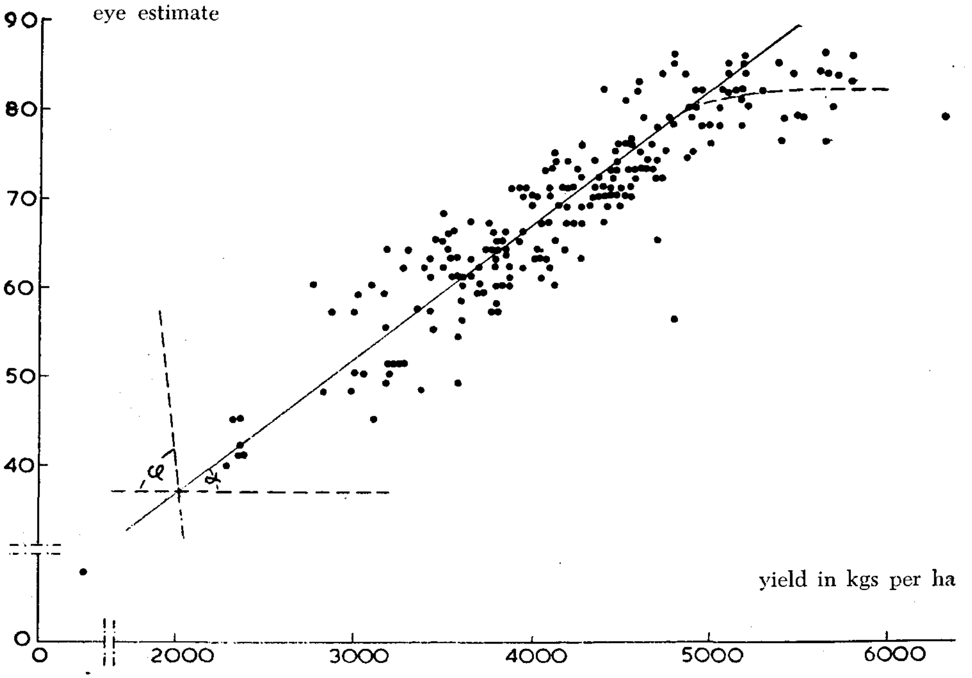

FIG. 2. RELATION BETWEEN THE EYE ESTIMATES OF THE YIELDS AND THE ACTUAL YIELDS OF OATS.

Service of Drainage, Land-improvement, and Re-allocation [1]). The scale runs from 1 to 100.

Figure 2 gives a picture of the result obtained in 1952. In contradistinction to 1951 the high yields were estimated too low in 1952; the high standard of yield in 1952 has ondoubtedly misled the valuers. These observations have disappeared in the calculations in order to obtain the necessary linear connection; the errors mentioned become thus somewhat too low.

The error of determining the yield by weighing was in accordance with the data of VISSER (1937) taken at 2% of the average yield. The average accidental error of estimate could be calculated from a large number of duplicate estimates.

Table 2 shows that there is a strikingly large difference between the estimates in 1951 and 1952. This difference, apart from what has been mentioned previously, will undoubtedly also have been caused by the fact that in 1952 much more care was given to the estimation.

As well as with grassland, also with oats there exists a systematic error of estimate which is about twice as large as the accidental error of estimate.

The question can be put, in what way is it possible to raise the accuracy of, for instance, the previously mentioned estimates to that of weighing. It looks attractive to presume that the increase of the number of independent repeats according to the square of the proportion of the total error of estimate

[1]) The estimates were effected in groups of 3 persons. Co-workers were: of the Research Department Messrs G. KORTOOMS, N. OOMKENS, G. H. REINDS, C. J. SCHOTHORST, P. STOFFEL, and of the Agricultural Experiment Station Messrs P. HOLSTEIN and R. H. SOL.

to the weighing error gives the required accuracy. It is however clear that an increase in the number of independent repeats decreases next to the accidental error only the systematic error on account of the subject and not that on account of the object. It depends on the proportion of the two last-mentioned as to what extent the accuracy of the estimates is increased by repeats.

The data about the oats estimates of 1952 enable us to form an opinion about this because in the execution a number of plots (24) were estimated each time in duplicate by different persons. A calculation shows that the share of the systematic error on account of the object is 233, of the systematic error on account of the subject is 98 kgs grain per ha.

This shows therefore that in this case a considerable part of the systematic error *cannot* be decreased by repeats. An improvement of the estimating technique will be essential to produce this effect.

We tender our thanks to the Director of the Central Institute of Agricultural Research and to the Head of the Research Department of the Government Service of Drainage, Land-improvement and Re-allocation for their collaboration in order to obtain this material.

LITERATURE

DEMING, W. E.: Statistical adjustment of data. London, 1946.

DRION, E. F.: Estimation of the parameters of a straight line and the variances of the variables, if they are both subject to error. Proc. Kon. Ned. Ak. 54 (1951) 256–260.

EZEKIEL, M.: Methods of correlation analysis. New York, 1950.

HEMELRIJK, J.: Over de bepaling van betrouwbaarheidsintervallen en schattingen van coëfficiënten van een rechte lijn uit een aantal onnauwkeurig waargenomen punten. Math. Centrum, Rapport S 28 (OL), 1952.

HOCKENSMITH, R. D.: Guide for soil conservations surveys. U.S. Dep. of Agriculture, Soil Conservation Science, Washington, 1948.

MURRAY, W. G.: Farm appraisal, classification and valuation of farm land and buildings. The Iowa State College Press, 1947.

POSTHUMUS, G.: Een mathematisch-statistische studie over de nauwkeurigheid van het organoleptisch onderzoek van boter. Thesis, University of Agriculture, Wageningen, 1943.

ROTHKEGEL, W.: Landwirtschaftliche Schätzungslehre. Stuttgart, 1952.

SCHEYGROND, W. en A. SONNEVELD: Vergelijking van verschillende typen Engels raaigras in een kunstweide onder beweidingsomstandigheden. Verslag van het Centraal Instituut voor Landbouwkundig Onderzoek over 1950 (1951) 50–54.

SITTIG, J.: Het probleem der waarnemingsfouten bij een massaal onderzoek. Statistica, 3 (1949) 49–68.

VISSER, W. C.: De ongelijkmatigheid van den grond en de nauwkeurigheid bij proefvelden. Versl. Landbouwk. Ond. No 43 (7) A, 1937.

VISSER, W. C. e.a.: Taxatie van gewassen. Landbouw 2, 1948.

VISSER, W. C.: Taxeren, een waardevolle steun bij onderzoekings- en voorlichtingswerk. Maandbl. Landbouwvoorlichtingsd. 8 (1951) 391–395.

WALD, A.: The fitting of straight lines if both variables are subject to error. Ann. Math. Statistics, 4 (1940) 284–300.

YATES, F.: Sampling methods for censuses and surveys. London, 1949.

Table 1. Results of the investigation into the accuracy of the eye estimates of the yields of grassland.

Year	Range of yields in kgs dry matter per ha	Mean yield in kgs dry matter per ha	Mean estimate	Mean error of weighing in kgs dry matter per ha	Unit of estimation equals ... kgs dry matter per ha	Total error of estimate in		Square of total error of estimate in % of variance of yields	Accidental error of estimate in kgs dry matter per ha	Systematic error of estimate in kgs dry matter per ha	Ratio error of estimate to error of weighing	Yield in kgs dry matter per ha corresponding with estimate = 0
						kgs dry matter per ha	% of range of yields (% of yield)					
1951	5300–15300	10200	42	573	160	992	9.9 (9.7)	19	< 512	> 848	1.7	3400
1952	3500–16200	10400	42	714	165	1250	9.5 (12.1)	25	< 528	> 1139	1.8	3400

Table 2. Results of the investigation into the accuracy of visual eye estimates of the yields of oats.

Year	Range of yields in kgs grain per ha	Mean yield in kgs grain per ha	Mean estimate	Mean error of weighing in kgs grain per ha	Unit of estimation equals ... kgs grain per ha	Total error of estimate in		Square of total error of estimate in % of variance of yields	Accidental error of estimate in kgs grain per ha	Systematic error of estimate in kgs grain per ha	Ratio error of estimate to error of weighing	Yield in kgs grain per ha corresponding with estimate = 0
						kgs grain per ha	% of range of yields (% of yield)					
1951	945–4682	3198	60	64	60	410	11.0 (12.8)	18	unknown	unknown	6.4	– 425
1952	251–6314	4166	68	83	67	278	4.6 (6.7)	15	125	253	3.3	– 450