Published July, 1971

Literature Cited

1. Brownstein, S. 1962. Energy barriers and structure of cis-, muco-, and allo-inositol hexaacetates in solution. Can. J. Chem. 40:870–874.
2. Eliel, Ernest L. 1969. Elements of stereochemistry. John Wiley & Sons, New York. 98 p.
3. Krzeminski, Z. S., and S. J. Angyal. 1962. Separation of inositols and their monomethyl ethers by gas chromatography. J. Chem. Soc. 3251–3252.
4. L'Annunziata, Michael F., and Wallace H. Fuller. 1971. Soil and plant relationships of inositol phosphate stereoisomers; the identification of D-chiro- and muco-inositol phosphates in a desert soil and plant system. Soil Sci. Soc. Amer. Proc. 35:000–000 (this issue).
5. Lemieux, R. U., R. K. Kullnig, H. J. Bernstein, and W. G. Schneider. 1957. Configurational effects on the proton magnetic resonance spectra of acetylated carbohydrates. J. Amer. Chem. Soc. 79:1005–1006.
6. Lemieux, R. U., R. K. Kullnig, H. J. Bernstein, and W. G. Schneider. 1958. Configurational effects on the proton magnetic resonance spectra of six-membered ring protons. J. Amer. Chem. Soc. 80:6098–6105.
7. Lichtenthaler, F. W., and P. Emig. 1968. NMR-studien an zuckern und cyclanolen V. Zusammenhang zwischen sterischer orientierung von O- und N-acetyl-gruppen und der signallage ihrer methyl resonanzen bei cyclit- und aminocyclit-polyacetaten. Carbohyd. Res. 7:121–137.
8. Suami, Tetsuo, and Kenji Yabe. 1965. Aminocyclitols. VII. A new synthetic route to myo-inosamine-4. Bull. Chem. Soc. Japan 38 (5):855–856.
9. Tate, M. E., and C. T. Bishop. 1962. Thin layer chromatography of carbohydrate acetates. Can. J. Chem. 40 (6): 1043–1048.

A SIMPLE STATISTICAL PROCEDURE FOR PARTITIONING SOIL TEST CORRELATION DATA INTO TWO CLASSES[1]

ROBERT B. CATE, JR. AND LARRY A. NELSON[2]

Abstract

Most soil test laboratories divide soil test results into two or more classes for the purpose of making fertilizer recommendations. This is usually done for the practical reason that it reduces the number of different recommendations necessary. However, the basis for defining the different classes (e.g., Very Low, Low, Medium, etc.) is often subjective or arbitrary. This paper explains a simple, yet statistically sound, method for setting the class limits. The procedure is to split the data into two groups, using successive tentative critical levels to ascertain that particular critical level which will maximize overall predictive ability (R^2), with the means of the two groups (classes) as the predictor values. The paper presents an actual example, using data believed typical of this kind of problem. Several continuous correlation models were also fitted to the same data. None gave as high an R^2 as a single Low-High split defined by the suggested procedure. Similar results have been obtained with a majority of 200 sets of soil test correlation data, indicating that the new procedure may be widely applicable.

Additional Key Words for Indexing:

SOIL TEST correlation data are generally characterized by considerable scatter. This is particularly true when soil test values are plotted against actual yields. To eliminate some of the scatter, most soil test correlation work uses relative or percentage yields. Relative yield is defined as being 100 times the yield of a treatment which provides adequate but not excessive amounts of all nutrients other than the one being correlated, divided by the yield of a treatment which is the same except that it includes the nutrient under study.

Although a single continuous model or function may be fit to soil test correlation data, many soil testing laboratories divide the results into just a few classes, such as Low, Medium, High, etc. This is usually done for the practical reason that it reduces the complexity of making recommendations. However, the basis for division into classes is often subjective or arbitrary. The purpose of this paper is to present a simple but statistically sound method for establishing the class limits.

It has previously been suggested that a dividing line between two categories (high probability of response and low probability of response) might be determined approximately by a graphical technique in which vertical and horizontal lines are superimposed on a scatter diagram so as to maximize the number of points in the positive quadrants.[3] The point where the vertical line intersects the X axis is used to divide the data into two classes. This dividing line has been termed the "critical level." The horizontal line has no real meaning, but it tends to fall about midway between the means of the classes. Handy tools for use in this graphical technique include either a transparent T-square or a plastic overlay marked with two perpendicular lines. In countries where the International Soil Fertility Evaluation and Improvement Program is active, this technique has been widely adopted to divide soils expected to give a relatively large response to a particular fertilizer nutrient from those expected to give little or no response, assuming that other nutrients are present in adequate amounts.

The graphical technique outlined above amounts to maximizing the computed chi-square value representing the test of the null hypothesis that the number of observations in each of the four cells (quadrants) is equal, i.e. the hypothesis that the data consist of one random population. However, since in practical interpretation it seems reasonable to use the means of the classes to predict average yield responses (in terms of percentages), a mathematical technique has been developed that is based directly on these means.

[1] Paper no. 3372 of the Journal Series of the North Carolina State Univ. Exp. Sta., Raleigh, N.C. This study was supported by the Agency for International Development, Department of State, Washington, D.C. Received Jan. 18, 1971. Approved Mar. 18, 1971.

[2] Research Associate, International Soil Fertility Evaluation & Improvement Program, Dept. of Soil Science, and Professor, Dept. of Statistics, North Carolina State Univ., Raleigh, N.C.

[3] Barr, A. J., and J. H. Goodnight. 1969. Statistical analysis system (SAS). Dept. of Statistics, North Carolina State Univ., Raleigh, N.C. NIH Project no. FR-00011.

The new technique consists of the following steps:

1) The data are ordered in an array based upon rankings of the X values i.e., soil test values. The (X, Y) pairs are maintained in this order throughout the analyses.

2) Starting with the X value that will place two or more points to the left of a vertical dividing line, one then calculates the corrected sums of squares of the deviations from the means of the two "populations" that result from moving to each successive X value. The sum of the two corrected sums of squares at each X level is then determined, and this pooled sum of squares is subtracted from the total corrected sum of squares of deviations from the overall mean of all Y observations. The difference between the pooled sum of squares and the total corrected sum of squares (the "between groups" sum of squares) is then expressed as a percentage of the total corrected sum of squares, or as R^2, since this difference represents the additional explanation obtained by fitting two means rather than one. It should be noted that the "between groups" sum of squares could be calculated directly by procedures commonly used in analysis of variance of one-way classification data.

3) By this simple iterative process, one obtains a series of R^2 values for divisions made at various levels of X. One picks the critical level of X as that where R^2 is maximum.

In other words, using this procedure one finds the value of X which best divides the data into two populations or classes, from the point of view of prediction. The method is general in the sense that an extension of the two-mean separation procedure may be used to divide the data into more than two populations. Probably the multi-class configuration which would be most useful would be a three-population model with the data being classified into high, medium, and low populations. Here the datum points would be separated into three populations by shifting the tentative class limits between low and medium and between medium and high in all possible ways. The separation chosen would be that for which a maximum "among-groups" sum of squares is obtained. Because there are two sets of class limits being shifted, the number of iterations required for the three-mean case would be somewhat greater than that required for the two-mean separation.

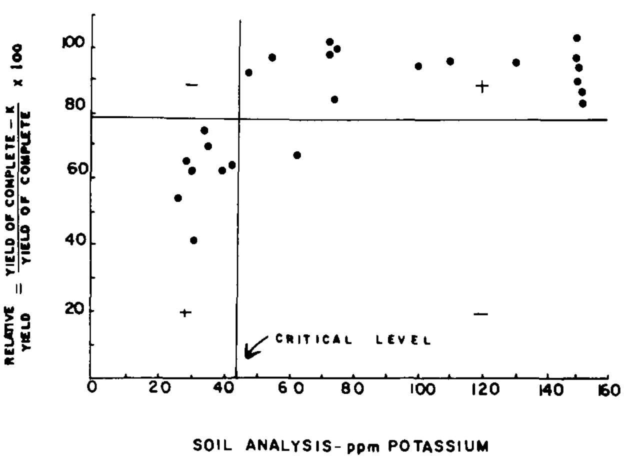

Fig. 1—Critical level by constructing cells to maximize computed chi-square for the null hypothesis that the overall distribution is random. Data are on correlation between soil K and plant response.

The old Cate-Nelson procedure[4] is illustrated in Fig. 1 and the calculations used in the new Cate-Nelson procedure are shown in Table 1. The original data for the example (Table 2) were taken from Freitas et al.[5]

The approach suggested here has been adapted by James Goodnight (Assistant Statistician, Dept. of Statistics, NCSU) for computer calculation using the Statistical Analysis System (SAS).[3] However, the technique is readily utilizable by anyone equipped with a desk calculator, or even with only a table of squares and a slide rule.

Several continuous models were also fitted to the same set of data. Equations for these models and results of the statistical analyses are shown in Table 3. It is interesting to note that none of these functions gave as high an R^2 as

Table 1—Calculations used in new Cate-Nelson procedure

Last value of soil potassium included in population 1	Mean relative yield in population 1	Corrected sum of squares of deviations from mean of population 1 (CSS-1)	Mean relative yield in population 2	Corrected sum of squares of deviations from mean of population 2 (CSS-2)	Postulated critical level (split between the two populations)	R-square for postulated critical level, i.e. TCSS minus sum of CSS-1 and CSS-2 divided by TCSS
28	59.5	60.5	85.5	6,301	29.0	0.16
30	60.7	68.7	86.6	5,769	30.5	0.23
31	55.8	358.8	88.9	3,584	32.5	0.48
34	59.6	655.2	89.6	3,380	34.5	0.47
35	61.3	745.3	90.7	2,974	37.5	0.51
40	61.6	747.7	92.4	2,160	42.0	0.62
44	61.9	752.9	94.1	1,306	46.0	0.73
49	65.4	1,670.	94.1	1,306	52.0	0.61
56	68.8	2,684.	93.8	1,280	62.0	0.48
68	68.6	2,686.	95.8	507.7	71.5	0.58
75	71.5	3,769.	95.2	452.2	76.0	0.44
77	73.5	4,369.	95.1	448.9	77.5	0.37
78	76.1	5,159.	95.6	310.2	90.0	0.28
102	77.4	5,570.	95.4	307.9	110.0	0.23
118	78.6	5,970.	95.0	300.0	126.0	0.18
133	79.6	6,290.	94.7	295.3	144.5	0.13

Totals: N = 24; Mean relative yield = 83.4; Total corrected sum of squares = 7,606 (TCSS).

Table 2—Percentage yield and soil test data employed in the study. For more information, see Freitas et al.[5]

Percentage yield	Soil test K, ppm
53.5	26
64.8	28
63.0	30
40.8	31
79.5	34
70.3	35
63.0	40
64.0	44
94.0	49
99.0	56
66.5	68
103.0	75
97.3	77
85.3	78
101.3	78
97.0	102
96.8	118
98.0	118
85.8	131
92.3	133
96.8	133
88.3	152
106.8	193
97.5	211

[4] Cate, R. B., Jr., and L. A. Nelson. 1965. A rapid method for correlation of soil test analyses with plant response data. Int. Soil Testing Series Tech. Bull. no. 1.

[5] Freitas, L. M. M. de, A. C. McClung, and F. Pimentel Gomes. May 1966. Determination of Potassium Deficient Areas for Cotton. Potash Review.

Table 3—Comparison of R^2 values obtained by fitting various models

Model	R^2	Equation for Model
Mitscherlich	.66	$\hat{Y} = A(1 - e^{-cX})$
Linear	.45	$\hat{Y} = b_0 + b_1 X$
Quadratic	.58	$\hat{Y} = b_0 + b_1 X + b_2 X^2$
Logarithmic	.59	$\hat{Y} = b_0 + b_1 (\log X)$
Reciprocal	.67	$\hat{Y} = b_0 + b_1 (\frac{1}{X})$
Cate-Nelson	.73	$\hat{Y} = b_0 + b_1 X$, where $\begin{cases} X = 0 \text{ if below critical level} \\ X = 1 \text{ if above critical level} \end{cases}$

the proposed procedure. Similar comparisons have been made on 200 sets of soil test correlation data. In well over half the cases, the two-population model give a higher R^2. Therefore, the method presented may have considerable utility.

USE OF SOIL CORES IN DETERMINING BULK DENSITY[1]

TYRONE M. GODDARD, E. C. A. RUNGE, AND W. M. WALKER[2]

Abstract

The bulk density of soil samples obtained with a hydraulic soil coring machine, for both core segments and clods, was compared to the bulk density of soil samples obtained with a Uhland sampler and natural clods taken from a soil pit. Subsurface, upper B, and lower B horizon samples were compared. Clods gave bulk density values that were significantly higher than cores in the upper solum, where small structural units occur; however, in the lower solum, where soil structural units are large or massive, variation between clods and cores was not statistically significant. There was no significant difference in bulk density due to the method of obtaining clods or cores.

***Additional Key Words for Indexing:* soil coring machines, volume weight determination.**

THE USE of soil coring machines for obtaining soil samples has increased rapidly in recent years. The volume of any core of known diameter and length can be calculated and is imminently suitable for bulk density determinations. This report compares the bulk density of six loessial soils, determined by natural clods and Uhland cores, with bulk density determined by clods and sections of duplicate soil cores obtained with a soil coring machine.

A discussion of methods for determining bulk density, except for the method using soil cores, was given by Blake (1). The method of Shaw (5), which was modified by Brasher et al. (2), measures the volume of a coated clod or ped of soil of known weight which is dipped in water for volume determination. The Coile method (3), from which the Uhland sampler was devised, measures the weight of the soil from a cylinder of known volume.

Materials and Methods

Four methods of measuring bulk density were compared. For convenience, they are identified as (i) coring core method, (ii) coring clod method, (iii) Uhland core method, and (iv) natural clod method. Soil coring tubes ranging in diameter from about 3–15 cm were available (Gidding Machine Co., Ft. Collins, Colo). Coring core samples in this study were obtained using a coring tube 7.6 cm in diameter by 120 cm long, as described by Runge (4). At appropriate depths, 7.6-cm core segments were taken from the soil core using a half cylinder with cutting edges on each end (Fig. 1). Samples were placed in plastic lined dispo-bags, and oven-dried at 105C for weight determinations. Sections of soil cores used for coring clod samples were obtained at identical soil depths from a duplicate soil core. Clod samples were placed in frozen food containers and taken to the laboratory. In the laboratory, samples were broken into clods and bulk density determined following the procedure of Brasher et al. (2). Uhland cores were obtained at the appropriate depths using a 7.6-cm diameter Uhland sampler with a 7.6-cm-long sleeve, following the procedure for determining bulk density outlined by Blake (1). Natural clods were obtained from a soil pit at the appropriate

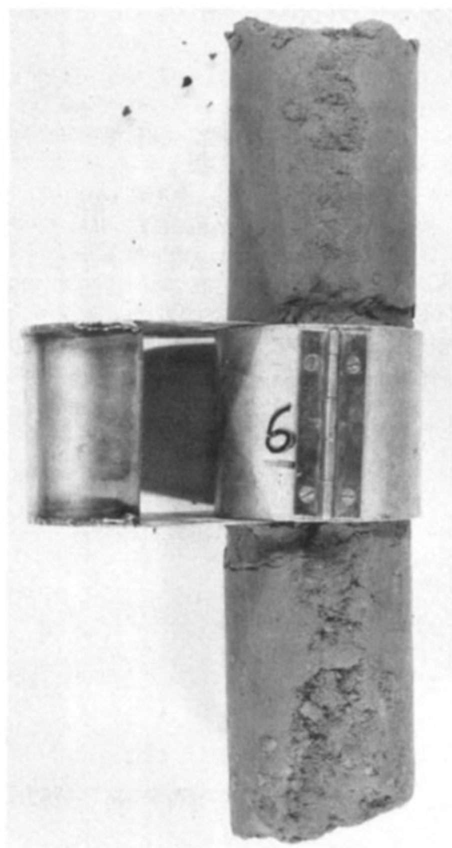

Fig. 1—Three inch core cutter used to obtain cores from hydraulic probe soil column. The inner cylinder is equipped with a hinge so that it may be fitted around soil core. The cylinder cover is equipped with cutting edges and is placed in position over inner cylinder for smooth cut.

[1] Contribution from the Dept. of Agronomy, Univ. of Illinois, Urbana-Champaign, 61801. A portion of this research was supported by funds from the Illinois Agr. Exp. Sta. Received Dec. 10, 1970. Approved Apr. 19, 1971.

[2] Graduate Assistant and Associate Professors, respectively.