1962

New techniques for data evaluation and control in soil engineering investigations

Coleman Anthony O'Flaherty
Iowa State University

Recommended Citation

O'Flaherty, Coleman Anthony, "New techniques for data evaluation and control in soil engineering investigations " (1962). *Retrospective Theses and Dissertations*. 2070.
https://lib.dr.iastate.edu/rtd/2070

O'FLAHERTY, Coleman Anthony, 1933–
NEW TECHNIQUES FOR DATA EVALUATION AND CONTROL IN SOIL ENGINEERING INVESTIGATIONS.

Iowa State University of Science and Technology
Ph.D., 1962
Engineering, civil

University Microfilms, Inc., Ann Arbor, Michigan

NEW TECHNIQUES FOR DATA EVALUATION AND CONTROL IN SOIL ENGINEERING INVESTIGATIONS

by

Coleman Anthony O'Flaherty

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Soil Engineering

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1962

TABLE OF CONTENTS

INTRODUCTION

The prosperity of any nation is bound up with the state of its roads. Nowhere is this more true than here in the United States. As the highways and byways of this country were improved and increased, vast new areas were opened to commercial use and it became possible to make use of the infinitely great natural resources of the virgin lands.

The growth of the road in the U. S. has not kept pace with other developments. The present highway system is taxed to its utmost and its future growth and improvement is regarded as a national necessity. All-weather roads, capable of carrying the modern day traffic for which they are designed, are required. A big drawback is the fact that all-weather roads are expensive to construct, especially in areas that lack suitable raw materials. As a result, many investigative efforts have been and are being undertaken in order to discover economical additives that will stabilize the in-place soil. The number of these investigated additives are numerous and the methods by which are evaluated are many.

Soil stabilization has not yet arrived at the stage whereby the results obtained by these accepted tests can be said to be "true" values. In many cases the design engineer is left with little knowledge as to what degree of confidence he can place in the results furnished to him. Too often the engineer overcomes this difficulty by specifying a high factor of safety in his highway design. It is believed that, in many instances, more credulity could be given to test results-and thus cause a lowering of the factor of safety-if statistical methods had been applied to the design and analysis of the experiment upon which the engineer's conclusions are based.

The application of statistical methods to engineering studies is not a new process as it has been going on for over thirty years. In 1924 W. A. Shewhart (1) wrote his paper "Some Applications of Statistical Methods to the Analysis of Physical and Engineering Data" for the Bell System Technical Journal and thus sowed the seed which led to the revolutionary growth in the application of statistical methods to industrial requirements. The decision to apply these methods to the engineering-cum-industrial fields was based on hard facts. The large industrial research laboratories and the great research institutions were interested in any techniques that would increase the return from the research and development dollar. They found that it was an economic-as well as scientific-fact that statistics, particularly the design and analysis of experiments, had a tremendous effect on the amount and quality of the information obtained from experimental work.

It is still, however, the belief of many researchers that statistical procedures are only useful at the end of an investigation-that is, when it comes to the final analysis of the data. This is not correct. Too often, after the data has been analyzed in this manner, the statistician has had to tell the investigator that his conclusions are, in fact, inconclusive: that they were bound to be so, as an inadequate amount of preparation went into the study and, as a result, inadequate data was obtained. The reasons for this are, of course, many-all depending on the type of work involved. Often times, however, much more reliable conclusions could be reached if a little more attention had been paid to the choosing of the experimental conditions, the number of tests to be run at the different experimental conditions, the number of samples per test condition or to the many other important variables involved.

Unfortunately, relatively little attention appears to have been paid to the application of statistical methods to soil stabilization studies. A search of the literature discloses many instances where certain amounts of statistics have been applied to the analyzing of data. However, in the great majority of these cases, the statistics appear to have been used only towards the end of the study and then only to suit particular purposes. Very many of the standard soil engineering tests are accepted with insufficient research into the variables involved in the tests. In rare instances are there statistical recommendations, specific to soil engineering studies, on how to analyze the data or to investigate the many variables involved.

It is felt that such studies and recommendations are needed in this new-yet old-science of soil engineering and, particularly, soil stabilization. It is towards helping to satisfy this need that this work was instigated.

PURPOSE OF THE INVESTIGATION

The primary purpose of this study was to provide some statistical methods which would help the soil engineering researcher to control and evaluate his results. In order to obtain data to illustrate and substantiate these recommendations, controlled experiments were conducted which involved two of the more common soil engineering tests. These experiments were also designed in order to give needed information from the purely engineering aspect.

In brief, this study was directed towards developing the following procedures-cum-recommendations:

1. Method for detecting outliers in a typical correlation study involving two methods of testing a stabilized soil.
2. Method for determining if a relationship, valid over a wide range of experimental conditions, exists between two methods of testing a stabilized soil.
3. Method for detecting outliers in a large series of soil-additive strength determinations involving triplicates of specimens.
4. Method for determining the reliability-as a whole-of an investigation that involves a large series of soil-additive strength determinations utilizing triplicates of specimens.
5. Method for detecting outliers in a large series of soil-additive strength determinations involving quadruples of specimens.
6. Recommendations regarding the number of specimens-per test condition-to use in an investigation involving large numbers of soil-additive strength determinations.
7. Method of selecting specimens for testing so as to minimize inherent specimen differences due to time or other factors.
8. Methods-involving the use of control specimens-to evaluate the validity of an investigation involving large numbers of soil-additive strength determinations.
9. Recommendations regarding the preparation of a soil sample prior to the actual investigation.
10. Method for detecting outliers in a series of soil-additive strength

determinations involving small numbers of strength determinations.

11. Method for evaluating the reliability of a curing chamber.
12. Method for determining if there is significant operator variability due to time trends.

To illustrate the above items, two investigations were conducted under controlled conditions. The first of these involved determining if a relationship, valid over a wide range of experimental conditions, existed between two methods of determining the strength of a stabilized soil. These two methods are the California Bearing Ratio and the unconfined compressive strength tests. The second of these controlled experiments involved determining the effects of fly ash and sodium carbonate as additives to soil-cement mixtures.

METHOD OF PRESENTATION OF DATA

This study involves statistical methods aimed toward soil stabilization investigations, examples of the use of such methods under actual experimental condtions and the engineering results obtained in the course of these experiments. As a result, there are many ways in which the data could be presented. One of the more obvious, perhaps, is that of a complete separation of the engineering and statistical works. It is felt, however, that the studies presented in this text are so interwoven that such a separation is not justified. For this reason, the data are presented in the following, most logical, manner.

The studies are presented in the sequence in which they actually took place and the analyses, both statistical and engineering, are shown in the same manner. The methods of test, materials involved and background information-both statistical and engineering-are also indicated where they can be most useful in interpreting the test results and understanding the procedures involved in the investigation.

The procedures-cum-results are presented in six parts, under the following titles:

Part 1 - Correlation Study Involving Two Methods for Testing Soils

This section contains details of items 1 and 2 as listed under "Purpose of the Investigation". The engineering results obtained in this phase of the main study are also reported here.

Part 2 - Detecting Outliers in a Large Series of Soil-Additive Strength Determinations

This part of the investigation contains details relating to items 3 and 4 as listed under "Purpose of the Investigation".

Part 3 - Further Methods for the Control of Data Quality

This phase of the investigation involves items 5 to 9 as listed under "Purpose of the Investigation". To obtain data to illustrate this part of the study, a controlled experiment involving the addition of fly ash and sodium carbonate to soil-cement mixtures was conducted.

Part 4 - Detecting Outliers in a Small Series of Soil-Additive Strength Determinations

This refers to item 10 as discussed under "Purpose of the Investigation".

Part 5 - Method for Evaluating the Reliability of a Curing Chamber and Operator Variability Due to Time Trends

This section contains details relating to items 11 and 12 as discussed under "Purpose of the Investigation".

It should be emphasized here that although the recommended procedures are presented under specific sub-titles, it is hoped and intended that they may also be utilized in other, but similar, types of soil engineering studies. For this reason, the recommended procedures are presented in as straightforward a manner as possible so that they may be used by the soil engineering researcher who, perhaps, is not too familiar with statistical terminology.

PRESENTATION OF RESULTS

Part 1 - Correlation Study Involving Two Methods of Testing Soils

One of the big difficulties in soil stabilization studies is the general inadequacy of existing testing procedures for determining the exact performance rating of an improved soil for highway pavement design purposes. New methods of testing are continually being devised to do this and generally attempts are made to correlate the results they give with those obtained by means of other known-to-be-reliable, although perhaps empirical, testing methods. At the moment, the most commonly used criterion for evaluating an improved soil is its unconfined compressive strength (UCS). The California Bearing Ratio (CBR) is also used as a design criterion for stabilized soils.

Purpose of the study

The objective of this phase of the investigation was, therefore, to determine whether a tight functional relationship, valid over a wide range of experimental conditions, existed between the two above methods of determining the strength of a cement-stabilized soil. It is hoped that the procedures presented here and the methods by which the data are analyzed will serve as prototypes for other future correlation work of a similar nature.

Background data

Although the UCS test is widely used in design, nevertheless the exact minimum design criteria have yet to be established. The British Road Research Laboratory suggests a minimum 7-day strength value of 250 psi for soil-cement, in order to withstand the requirements of the American Society for Testing Materials (ASTM) durability test for wetting and drying or freezing and thawing (2). The Portland Cement Association states that soil-cement having an UCS of 300 psi after 7 days will usually pass the durability tests (3).

The California Bearing Ratio test is very much used in flexible pavement design. An excellent description of the history and development of the test is given in the 1950 Transactions of the American Society of Civil Engineers (4). It is known that at least 14 of the 50 states now use the CBR value of a soil as their principal strength standard in highway design (5). The CBR test has been extensively correlated with the field performance of soils and it has been found that materials directly under the bituminous surface of a highway should have a CBR of at least 80% -this is equivalent to a laboratory CBR of about 120% (6). Lower CBR values are allowed at greater depths as the wheel-load stresses are more widely distributed.

While the reliability of the CBR test for pavement design purposes is excellent, nevertheless the test has many disadvantages. Not only does the test require large quantities of soil and stabilization materials, but, in addition, it is relatively difficult and time-consuming to perform. The penetration test itself requires the services of two people for a minimum of ten minutes actual penetration time. On the other hand, the UCS test is simple to perform and requires small volumes of soil. To illustrate, it may be mentioned that in the work described here, each CBR specimen took at least one hour to prepare and test, whereas each pair of UCS samples required, on an average, a maximum of fifteen minutes to prepare and test. The amount of soil required for each CBR test was about ten pounds while only about one half-pound was required for each UCS test.

Materials

The soil used in this investigation was a dune sand typical of those found in eastern Iowa. Sampling location and properties of the soil sample are given in Table 1.

Table 1. Description and properties of sand

Location:	Benton County, Western Iowa Sect NW 1/4, SE 1/4, S-16 Twp 86 N, Rn 10 W	Geological description	Wisconsin-age eolian sand, fine-grained, oxidized, leached
Soil series:	Carrington	Horizon:	C
Sampling depth, in feet:	6 - 11	IEES code number:	S-6-2
Textural composition, %:[a]		Mineral composition, %:[b]	
Gravel (2 mm.)	0.0	Total quartz	73.4
Sand (2 - 0.074 mm.)	94.4	Total feldspar	19.9
Silt (74 - 5μ)	1.6	Rock fragments	3.2
Clay ($<5\mu$)	4.0	Calcite	0.2
Colloids ($<1\mu$)	3.5	Mica	Trace
		Total heavy minerals	1.0
		Minus 0.044 mm. material	2.5
Predominant clay material:[c]	Montmorillonite and illite interlayer	Physical properties:	
		Liquid limit, %	9.0
Specific gravity 25C/4C:	2.64	Plastic limit, %	-
		Plasticity index	Non-plastic
Chemical properties:		Classification:	
Carbonates, %[d]	0.02	Textural	sand
pH	6.5	Engineering (A.A.S.O.)	A-3(0)
Organic matter, %[d]	0.04		

---

[a] Dispersed by air-jet with sodium metaphosphate dispersing agent. Coarse sand, 12.9%; fine sand, 81.5%.

[b] Material larger than 0.044 mm. (Per cent by volume of the whole sample)

[c] From X-ray analysis.

[d] Per cent by weight of oven-dry soil.

The cement used was a Type I normal Portland cement. Its properties are given in Table 2.

Table 2. Cement properties[a]

Cement type: I		
Chemical composition, %:		
Silica	21.62	
Alumina	5.04	
Iron oxide	2.97	
Lime	64.05	
Magnesia	2.90	
Sulfur trioxide	2.26	
Ignition loss	0.58	
Insoluble residue	0.16	
Physical properties:		
Fineness, turbidometer (Wagner), sq. cm. /gm.		1855
Fineness, air permeability (Blaine), sq. cm. /gm.		3395
Compressive strength (1:2.75 mortar)		
1 day, psi		---
3 day, psi		2269
7 day, psi		3721
28 day, psi		5625

[a] Data supplied by Penn-Dixie Cement Corporation, Des Moines, Ia.

Experimental procedure

Soil sample The first step consisted in preparing a large homogenous master batch of the sandy soil from which 156 sub-batches were randomly selected. These were then partitioned into 39 quadruple sub-batches. Members of the same quadruple were then handled in the same manner, the same specified amounts of cement and water being added to each. Each such quadruple member was then sub-divided into three specimens for UCS testing.

The experimental origins of the CBR values and the UCS values are perhaps best illustrated schematically as in Figure 1.

CBR test All CBR specimens were prepared and tested according to the ASTM "Tentative Method of Test For Determining the Bearing Ratio (CBR) of Soils, 1959", with some exceptions as described now.

Due to the large number of CBR specimens that had to be prepared and tested, it was felt to be impractical to use the CBR mold as described by ASTM. Instead a special CBR mold was devised. This mold is shown in Figure 2. It consists, simply, of a standard CBR mold cut on one side only. A 1/16 in. wide piece of steel, of such size and contour as to replace exactly the milled material, was then inserted into the gap and soldered onto one side of the mold. The gap was then closed or opened as required by means of the bolt attachment shown in Figure 2.

Utilizing this mold, CBR specimens were prepared in the following manner. Using the bolt attachment, the gap on the side of the mold was closed as tightly as possible using a hand wrench. The inside of the mold was then lightly coated with oil. The mold-with collar attached-was clamped to the base plate and the spacer disk inserted into the mold. Two circular layers of wax paper, each just under 6 in. in diameter, were placed on top of the disk. The soil, cement and water mixture was then compacted in the mold in accordance with the standard procedure (7). After compaction the extension collar was removed and the compacted material was carefully trimmed so as to be even with the top of the mold. The spacer disk and base plate were then removed and the mold plus compacted material weighed. A piece of wax paper, approximately 7 1/2 in. square, was placed on each end of the mold and fastened in place by means of tight elastic bands. The mold was then placed in the curing room for 24 ± 3 hours. Care was taken that the mold rested on the end at which the soil cement was trimmed level with the lip of the mold. After this curing period, a mark was made on the mold lip and a similar adjacent mark was made on the soil-cement specimen. The bolt attachment on the outside of the mold was then loosened, allowing the mold to open about one-fourth of an inch. Usually this was sufficient to allow the mold to be withdrawn from the specimen. The specimen was then carefully wrapped in wax paper, sealed with adhesive tape, and then replaced in its original position in the curing room.

In order to perform the penetration test, the CBR specimen was unwrapped and replaced in its original mold so as to fit its original contour. This was checked by having a mark on the specimen line-up

Figure 1. Structure of the experiment involving the California Bearing Ratio and unconfined compressive strength tests.

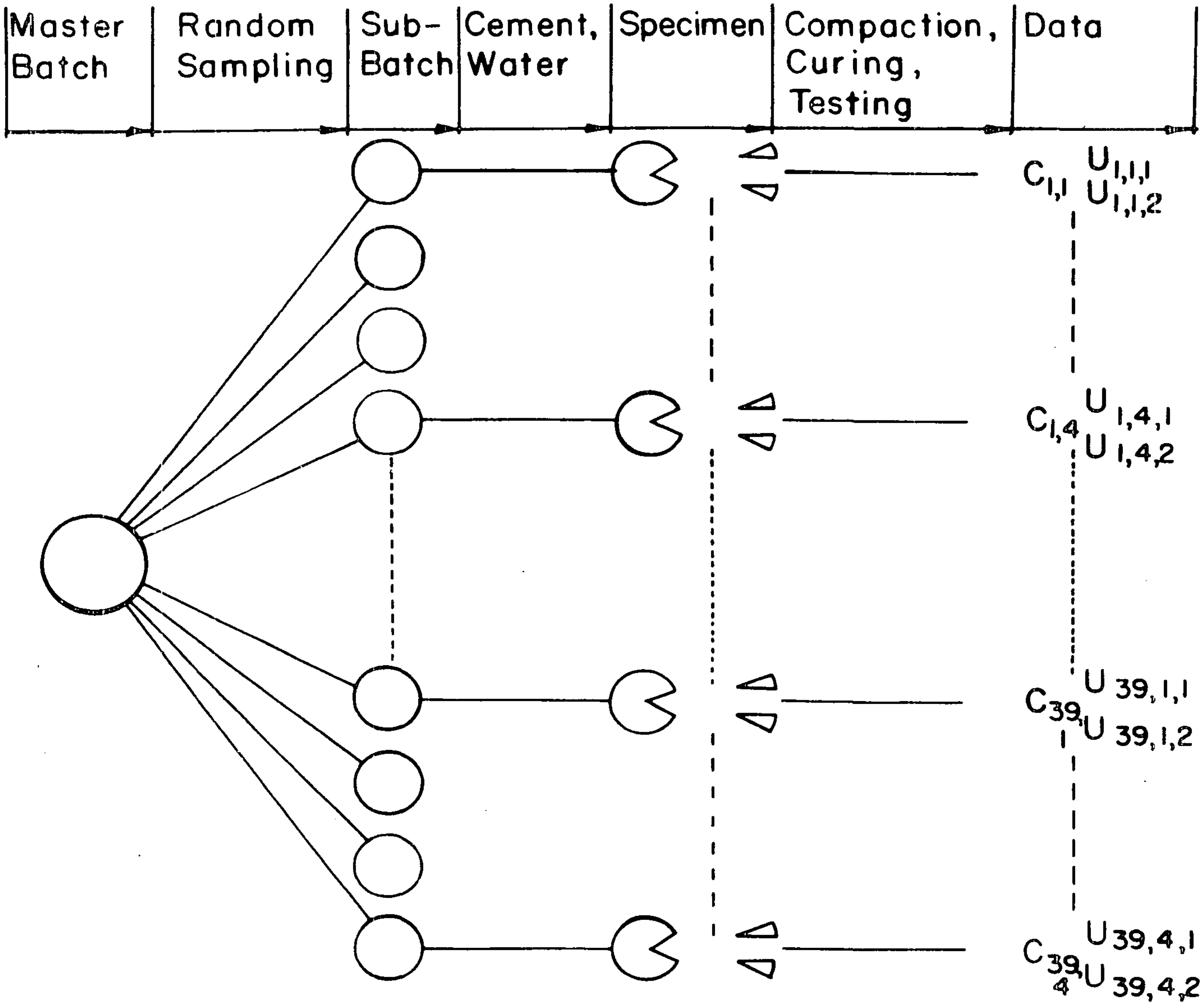
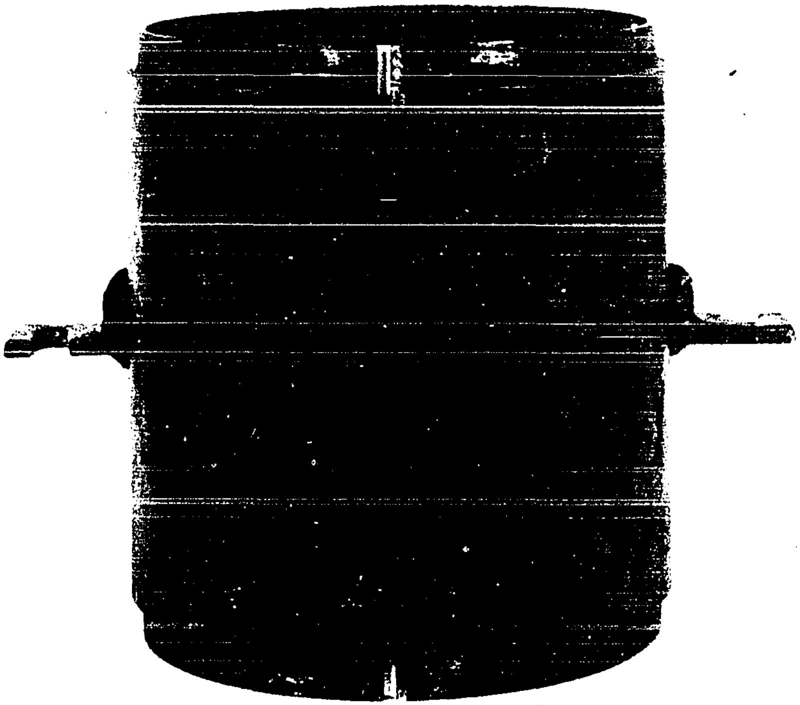

Master Batch
Random Sampling
Sub-Batch
Cement, Water
Specimen
Compaction, Curing, Testing
Data
$C_{1,1}$
$U_{1,1,1}$
$U_{1,1,2}$
$C_{1,4}$
$U_{1,4,1}$
$U_{1,4,2}$
$C_{39,1}$
$U_{39,1,1}$
$U_{39,1,2}$
$C_{39,4}$
$U_{39,4,1}$
$U_{39,4,2}$

Figure 2. Modified California Bearing Ratio mold.



with a similar mark on the mold. The mold was then closed about the specimen using a hand wrench to tighten the bolt attachment as tightly as possible. The penetration test was then immediately carried out upon the specimen and the CBR value read at 0.10 inch penetration.

Some specimens were soaked before being tested. In such cases, the CBR specimens were taken from the curing room, unwrapped and immersed in distilled water. Care was taken that the water surface remained a constant 1 1/2 in. above the top of each specimen. The soaking period for such specimens was 24 ± 2 hours.

UCS test Specimens used in this test were 2 in. in diameter by 2 in. high. They were molded and compacted using a drop-hammer molding apparatus developed by Davidson and Chu (8). After compaction, specimens were ejected from the molds with a hydraulic jack. Each specimen was weighed to the nearest 0.1 g. and its height measured to the nearest 0.001 in. A height tolerance of ± 0.05 in. was maintained on all specimens.

Each specimen was wrapped in wax paper and sealed with adhesive tape before being placed in the curing room. After curing, the unconfined compressive strength of each specimen was obtained by means of a testing machine of the proving ring type. Load was applied to each specimen, the rate of deformation being 0.10 in. per minute, until complete failure was reached. The maximum applied load in pounds was divided by the cross-sectional area of the 2 in. diameter specimen and the result, in psi, reported as the unconfined compressive strength of the specimen.

Certain 2 in. diameter by 2 in. specimens required soaking prior to testing. Such specimens were unwrapped and immersed in distilled water for 24 ± 2 hours. Care was taken that, at all times, the surface of the water was one-fourth of an inch above the top of each sample.

Dry Densities One CBR specimen and two UCS specimens were prepared from each sub-batch. A moisture sample was taken immediately prior to the preparation of the first specimen and immediately after the compaction of the last specimen. The average of these two moisture contents was then used to calculate the dry densities of the three specimens prepared from that particular sub-batch. The dry densities of the two types of specimens were judged to be within acceptable limits of variation. The average dry densities obtained at varying cement contents are shown in Table 3.

Table 3. Average dry densities of soil-cement specimens in pounds per cubic foot

Type of specimen	Cement content, % of total mix		
	5	8	11
UCS	107.2	111.0	112.1
CBR	105.0	108.7	110.8

Curing Each CBR specimen and its corresponding pair of UCS specimens were placed side by side in the curing room. The temperature in the curing room was maintained at 70° F. and the humidity at 90% relative humidity. Moist curing periods varied from 10 hours to 21 days. Approximately half of the specimens were then cured for a further 24 ± 2 hours by immersing them in distilled water.

Cement contents Specimens were prepared using cement contents of 5, 8 and 11 percent by weight of total mix.

Statistical analysis of data

Such a large investigation as this almost always reveals several suspect quadruples which, although they "fall out of line" in some respect, cannot be eliminated by pointing to known and noticed causes. Some criteria are therefore needed whereby suspect values can be deemed either "true" or "false". The following procedure for detecting outliers is believed to be applicable to investigations such as this.

Outlier analysis The data is examined for inhomogeneity in two different respects. The point of view adopted here is that a particular quadruple should be discarded if it appears to be suspected in both examinations.

Figure 3. Unconfined compressive strength-time relationship for soil-cement mixtures.

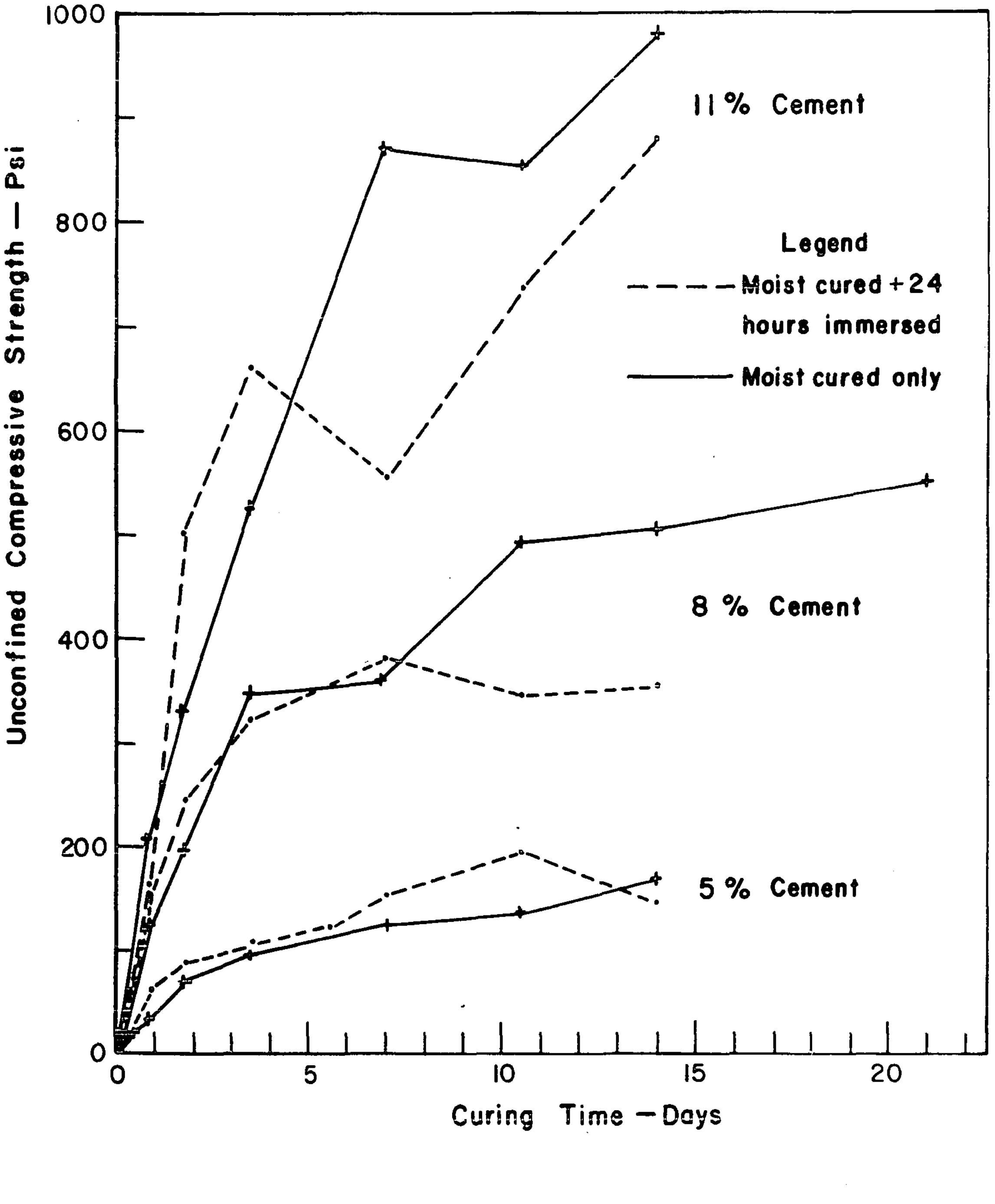
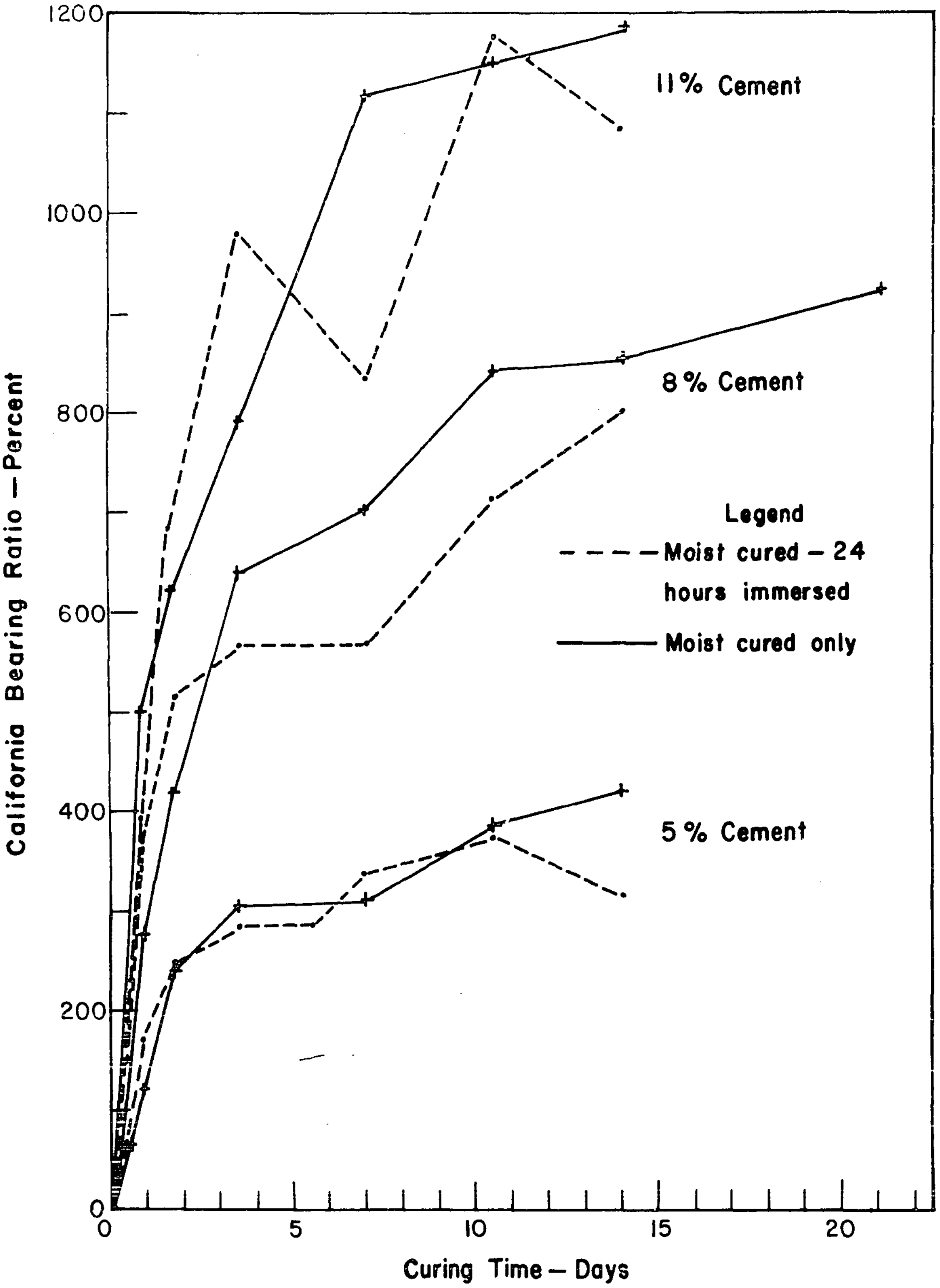

1000
800
600
400
200
0
Unconfined Compressive Strength — Psi
0
5
10
15
20
Curing Time — Days
11 % Cement
8 % Cement
5 % Cement
Legend
Moist cured + 24 hours immersed
Moist cured only

Figure 4. California Bearing Ratio-time relationship for soil-cement mixtures.


1200
1000
800
600
400
200
0
California Bearing Ratio — Percent
0
5
10
15
20
Curing Time — Days
11% Cement
8 % Cement
5 % Cement
Legend
Moist cured — 24 hours immersed
Moist cured only

Deviations of quadruple averages from trends exhibited in Figures 3 and 4   Figures 3 and 4 indicate the strength results obtained under varying conditions. The values upon which these graphs are based are shown in Appendix A. Examination of these graphs indicates that, perhaps, some values should be suspected. The most obvious are the 7-day strength values-both UCS and CBR-for the 11 percent cement, moist cured specimens. Another, perhaps, is the 14 day, immersed CBR value obtained with 11 percent cement.

Unusual consistency in quadruple configurations   This test is intended to give a look at internal quadruple structure.

As mentioned already, the first step consisted in preparing a large homogeneous master batch from which 152 sub-batches were randomly selected. These were then partitioned into 38 quadruple sub-batches. Members of the same quadruple were then handled in the same manner, the same specified amount of cement and water being added to each. Each such member was then sub-divided into three specimens, one large specimen for the CBR testing and two smaller specimens for the UCS testing. These three specimens were then cured in the same manner, for a specified length of time, before being tested.

For convenience, the 456 resultant strength values are labelled C_{ij} and U_{ijk}, where i varies from 1 to 38, j from 1 to 4 and k from 1 to 2. Thus, as indicated in Figure 1, C_{ij} denotes the CBR figure from the jth sub-batch at the ith set of factor levels and U_{ijk} denotes the UCS figure for the kth aliquot (equal part) prepared from the jth sub-batch at the ith set of factor levels.

Let $X_1 \leq X_2 \leq X_3 \leq X_4$ be the four strength values, ordered from lowest to highest, obtained from each quadruple set. It is now necessary to examine the differences between X_1 and X_2, X_2 and X_3, and X_3 and X_4 as the proposed statistic will involve these gaps. It is to be expected-even under the assumption of homogeneity-that both of the gaps $(X_4 - X_3)$ and $(X_2 - X_1)$ will tend to be larger than the gap $(X_3 - X_2)$. It is therefore necessary to adjust them to equal expectation. This is most easily done by multiplying $(X_3 - X_2)$ by

$$a = E(X_2-X_1)/E(X_3-X_2) = E(X_4-X_3)/E(X_3-X_2) = 1.2329$$

as indicated in reference (9).

The proposed outlier test procedure now requires the computation of the following statistic:

$$R_o = \frac{\text{largest of } (X_3-X_2),\ a(X_3-X_2),\ (X_2-X_1)}{\text{2nd largest of } (X_4-X_3), a(X_3-X_2), (X_2-X_1)}$$

for each set of quadruple determinations. The purpose for so doing is in order that the sample cumulative distribution function (CDF) of the R_o values can then be examined with respect to the maximum absolute deviation, D_n, between this sample CDF and the theoretical CDF of R_o as obtained under the normality assumption. The hypothesis that the entire series is homogenous is rejected if this statistic is too large.

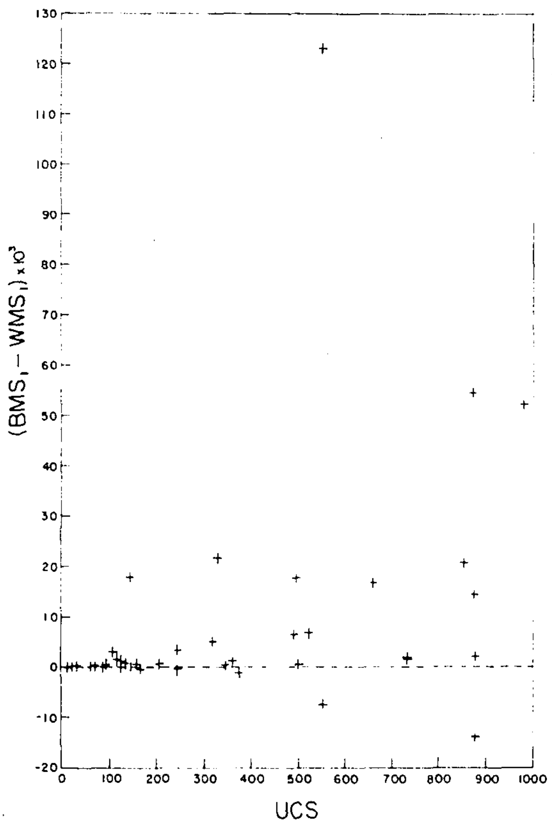
Before actually computing for this statistic it is necessary to consider the following two factors.

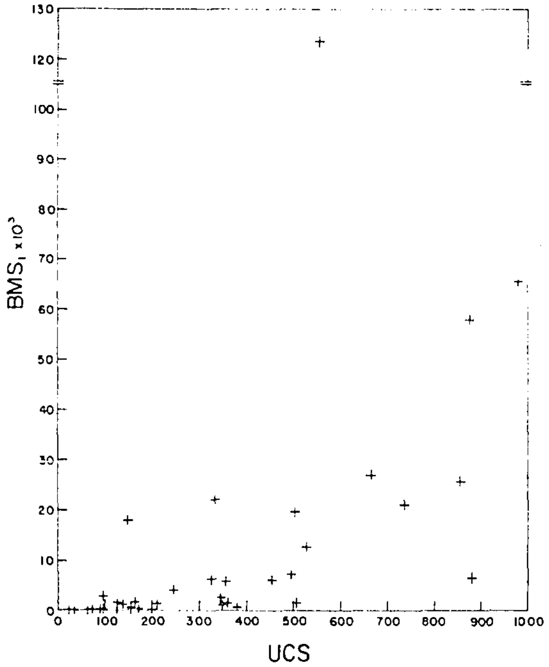
A. Should the adjusted gap ratios R_o be computed for quadruples of UCS aliquot averages $\bar{U}_{ij}$; j = 1, 2, 3, 4, or should they be computed for octuples U_{ijk}; j = 1, 2, 3, 4; k = 1, 2?

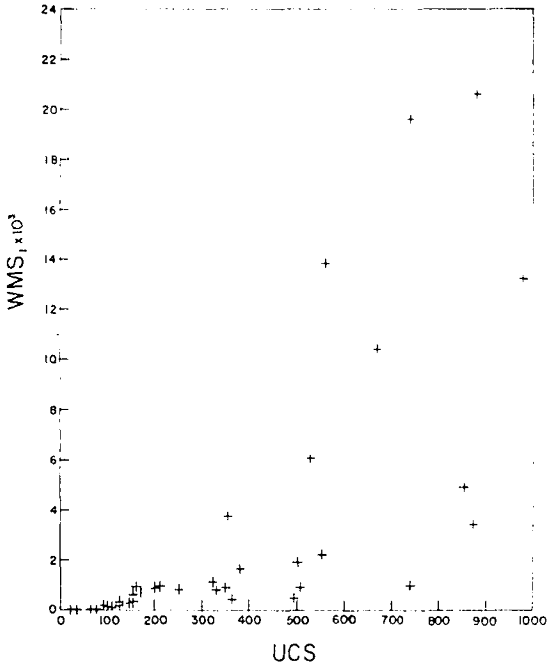
B. If the R_o statistics are computed for UCS quadruples U_{ij}, should their R_o's be pooled with the R_o's computed for the quadruple CBR determinations in one over-all test involving all 76 R_o values or, alternatively, should separate tests be conducted, one involving only the 38 UCS values, the other involving the 38 CBR values?

The answers to these questions depend upon the correlation structure of the data. If the UCS data show no intra-sub-batch correlation, then either a quadruple or octuple approach to the UCS data is correct, with preference likely given to the more informative octuple approach. On the other hand-unless special distributions are computed-only the quadruple approach is possible if intra-sub-batch correlation does exist. As for the second question, either approach is valid if the R_o's for the UCS series are independent of the R_o's for

Figure 5. Assessment of the relationship between inter-sub-batch variance and batch mean.


(BMS1 – WMS1) x10³
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
-20
0
100
200
300
400
500
600
700
800
900
1000
UCS


BMS1 x10³
130
120
110
100
90
80
70
60
50
40
30
20
10
0
0
100
200
300
400
500
600
700
800
900
1000
UCS

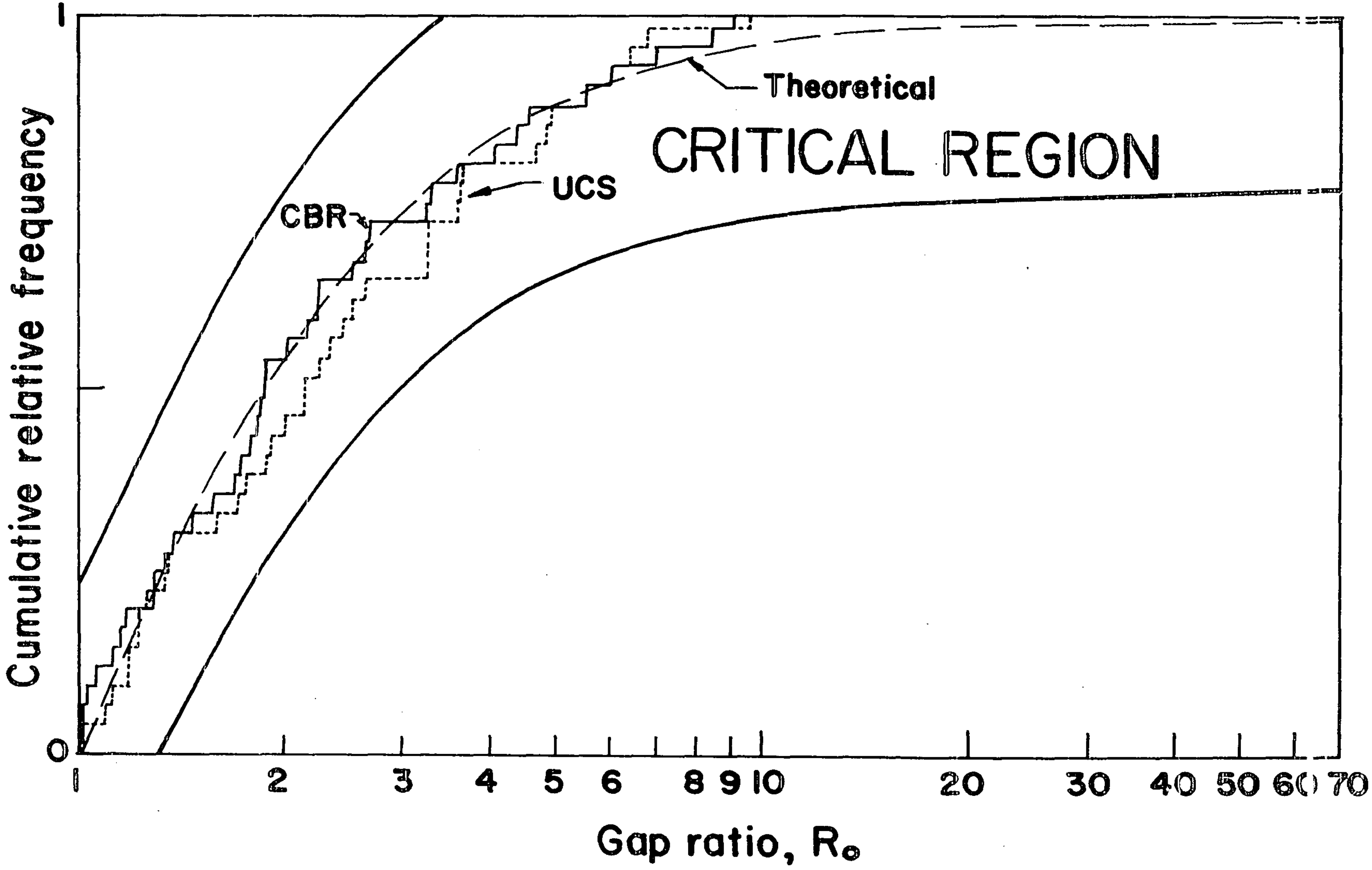

WMS1 x10³
24
22
20
18
16
14
12
10
8
6
4
2
0
0
100
200
300
400
500
600
700
800
900
1000
UCS

the CBR series. The pooled approach is to be preferred if it is desired to have "experiment-wise" error control; separate tests are to be preferred if it is desirable to examine the two series separately, as for instance if two operators are involved in the two series.

It is easily verified that intra-sub-batch correlation does exist for the UCS series. Let WMS_i be the four-degree-of-freedom within-sub-batch (or between-aliquot) mean square, and BMS_i the three-degree-of-freedom between-sub-batch mean square computable on the basis of the octuple $(U_{i11} \ldots U_{i42})$. The 38 points $(\bar{U}_{i..}, BMS_i - WMS_i)$ are plotted as indicated in Figure 5. These 38 points, far from hovering about zero, tend to fall on an upward turning parabola whose vertex is at the origin, thus establishing that there is a within-sub-batch correlation. This is corroborated by the plots of Figures 5a and 5b. Intra-sub-batch correlation is, of course, to be expected, in view of the structure of the experiment as outlined in Figure 1. At any rate, its presence eliminates the possibility of conducting the UCS analysis on the basis of octuples.

It remains only to check on the dependence, if any, of the 38 R_o's computed for quadruples of UCS aliquot averages $\bar{U}_{ij}$ and the 38 R_o's computed for quadruples C_{ij} of single CBR determinations. Here, the existence of the UCS intra-sub-batch correlation, which of course amounts to the existence of a UCS sub-batch effect, makes it plausible that some such dependence exists, since, as indicated by Figure 1, a given sub-batch will yield both a member C_{ij} of a CBR quadruple and a member $\bar{U}_{ij}$ of a UCS quadruple. This plausibility is reinforced by certain additional features of the experiment not exhibited by Figure 1 such as the fact that the single CBR specimen and the two UCS specimens (or aliquots) made from the same sub-batch were cured side by side in the curing chamber. It is therefore surprising that the data nevertheless indicate that the two sets of R_o's are, in fact, independent. This was verified by transforming all 76 R_o's to unit-normality and computing the coefficient of correlation $r = -.016$, which, being negative, requires no further computation for acceptance of independence. This lack of dependence now makes it possible to choose between conducting two

Figure 6. Theoretical and empirical CDF's for the UCS and CBR data: critical region is for a 2 1/2% test


Cumulative relative frequency
0
1
CBR
UCS
Theoretical
CRITICAL REGION
1
2
3
4
5
6
8 9 10
20
30
40 50 60 70
Gap ratio, Ro

separate tests and conducting a single pooled test. Since two different operators actually were involved in the CBR and UCS series, it was decided to forego the advantage of experiment-wise error control, and to conduct two separate tests. This step turned out to be rewarding, since the suspected greater reliability of the operator for the CBR series seems borne out by the better behaviour of the empirical CBR CDF as indicated in Figures 6 and 7.

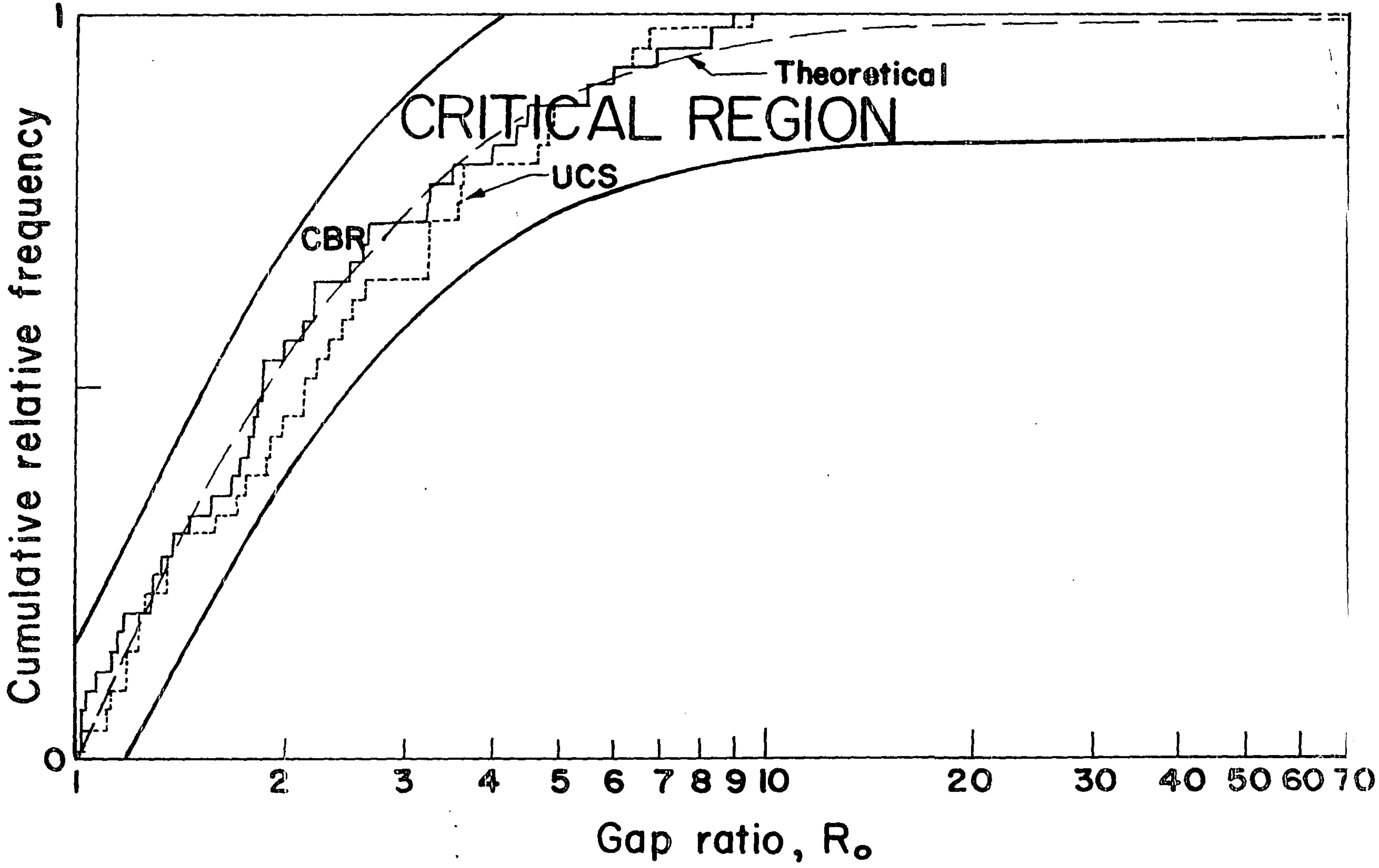
In Figures 6 and 7 are plotted the theoretical CDF and the sample CDF's for the R_o values obtained from the UCS and CBR quadruples. The critical region in each figure is in fact a critical region for the sample CDF of R_o's, if this sample CDF is considered so that the portions at height 0 and 1 are deleted. Two critical regions are given. The critical region in Figure 6 is for a 2 1/2 percent test and consists of comparing D_n to a constant ϵ that is exceeded by D_n with probability only $1-.95^{1/2}$ when all 38 quadruples of the series are internally homogeneous. This constant ϵ is computed from Millers formula (10) where $n = 38$, $a = (1-.95^{1/2})/2$ and $A(a) = .17$. This gives a value of $\epsilon = .2347$, which means that the total vertical distance between the curved critical-region boundaries is .4674.

It could perhaps be argued that, in a large investigation such as this, a 2 1/2 percent test is too stringent and unrealistic. The critical region corresponding to a level of .30 is shown in Figure 7. ϵ then becomes .55.

It is clear that in this investigation all the strength values appear to be statistically valid.

Detectable inhomogeneities A final remark concerns the types of inhomogeneities that will be detected by the above test procedure. Any feature of the series leading to undue accumulation at specific R_o values-such as an operator fabricating determinations or split-plotting-will be detected. In addition, isolated "single splits" among population means, of form (μ——μ-μ-μ), (μ-μ——μ-μ) or (μ-μ-μ——μ) will be detected. "Double splits" and "triple splits" of type (μ-μ——μ——μ), (μ——μ-μ——μ), (μ——μ——μ-μ) and (μ——μ——μ——μ) will not be

Figure 7. Theoretical and empirical CDF's for the UCS and CBR data: critical region is for a 30% test


Theoretical
CRITICAL REGION
UCS
CBR
Cumulative relative frequency
1
0
1
2
3
4
5
6
7
8
9
10
20
30
40
50
60
70
Gap ratio, R_o

detectable. Parenthetically, this lack of power against double splits is the penalty paid by R_o for good power against single splits.

Correlation study There are many procedures available in statistical methodology whereby it can be determined if a relationship exists between two or more variables. These standard procedures are most useful when either none or only one variable is subject to error and when it is required to predict the dependent variable from the known independent variable.

Standard procedures, however, are not available for the problem presented in this experiment since both the UCS and the CBR measurements are subject to error. The following procedure, therefore, is presented as being a logical approach to a problem such as this. The primary purpose is to determine if there is, in fact, a direct relationship between the results obtained using both methods of testing and, in so doing, to determine what is the best line-of-fit that will most nearly minimize errors in prediction.

Regression analysis The first step in the analysis was to plot all the CBR values against their corresponding UCS values, using linear graph paper. This plot is shown in Figure 8. This grouping of the data strongly suggested that a relationship does exist between the results obtained from both methods of test. When the data was transformed by means of logarithims, as shown in Figure 9, it appeared very likely that this relationship might be a linear one between log CBR and log UCS.

Figures 9, 10 and 11 now suggest that the following statistical model (*) will afford a reasonable description of the data if it is hypothesized that a simple functional relationship relates UCS logs to CBR logs in the absence of test errors.

$$X_{ij} = \log(\mathrm{UCS}_{ij}) = \xi + \epsilon_{ij}$$

$$Y_{ij} = \log(\mathrm{CBR}_{ij}) = \alpha + \beta\xi_i + \eta_{ij}, \quad i{:}1, 2. \ldots 39, j{:}1, 2, 3, 4$$

where

ξ_i = errorless log UCS for the i^{th} test condition

Figure 8. Relationship between CBR values and UCS values for soil-cement mixtures.

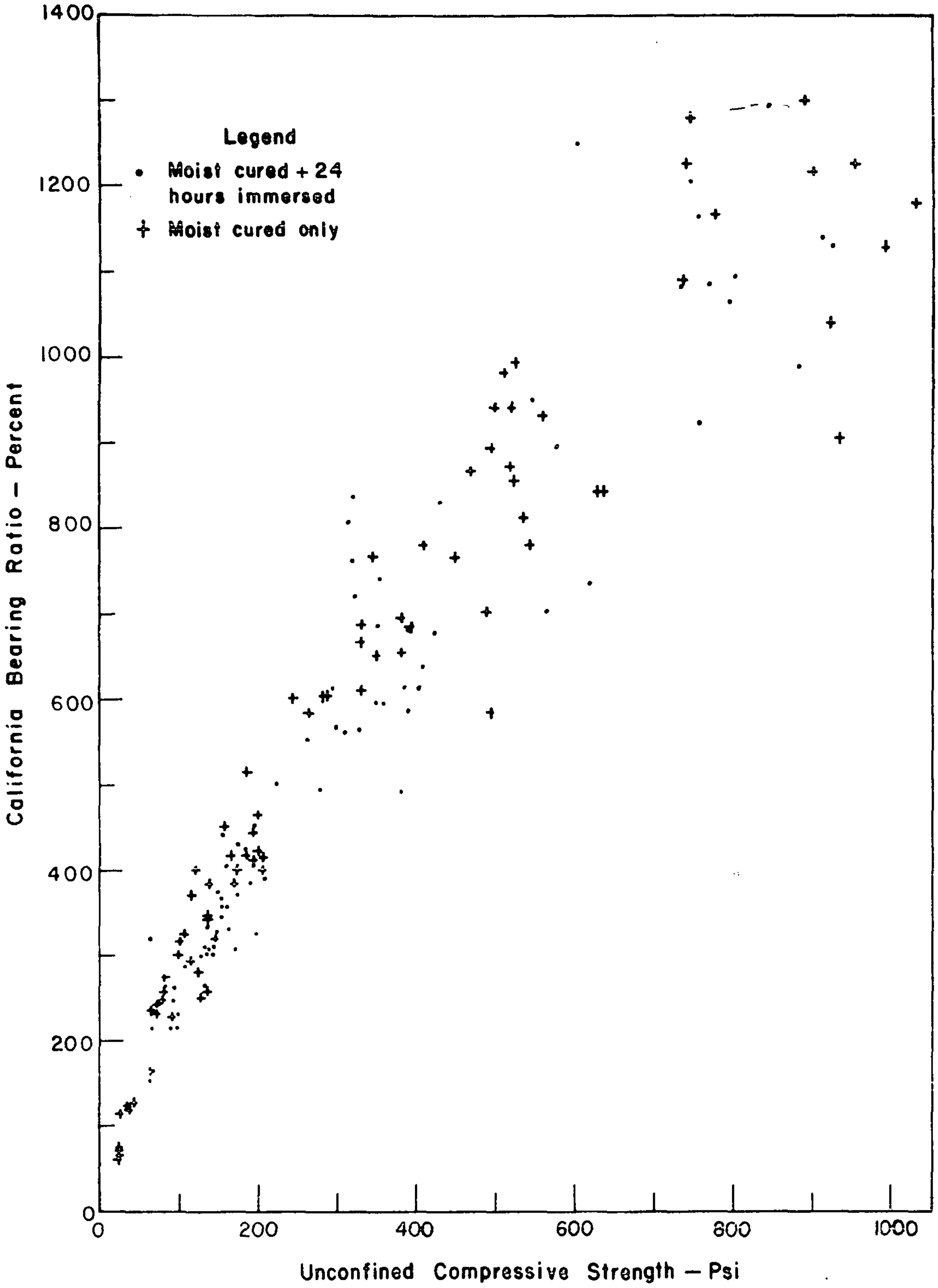

Legend
• Moist cured + 24 hours immersed
+ Moist cured only
California Bearing Ratio – Percent
Unconfined Compressive Strength – Psi
0
200
400
600
800
1000
1200
1400

Figure 9. Relationship between log CBR and log UCS values for soil-cement mixtures.

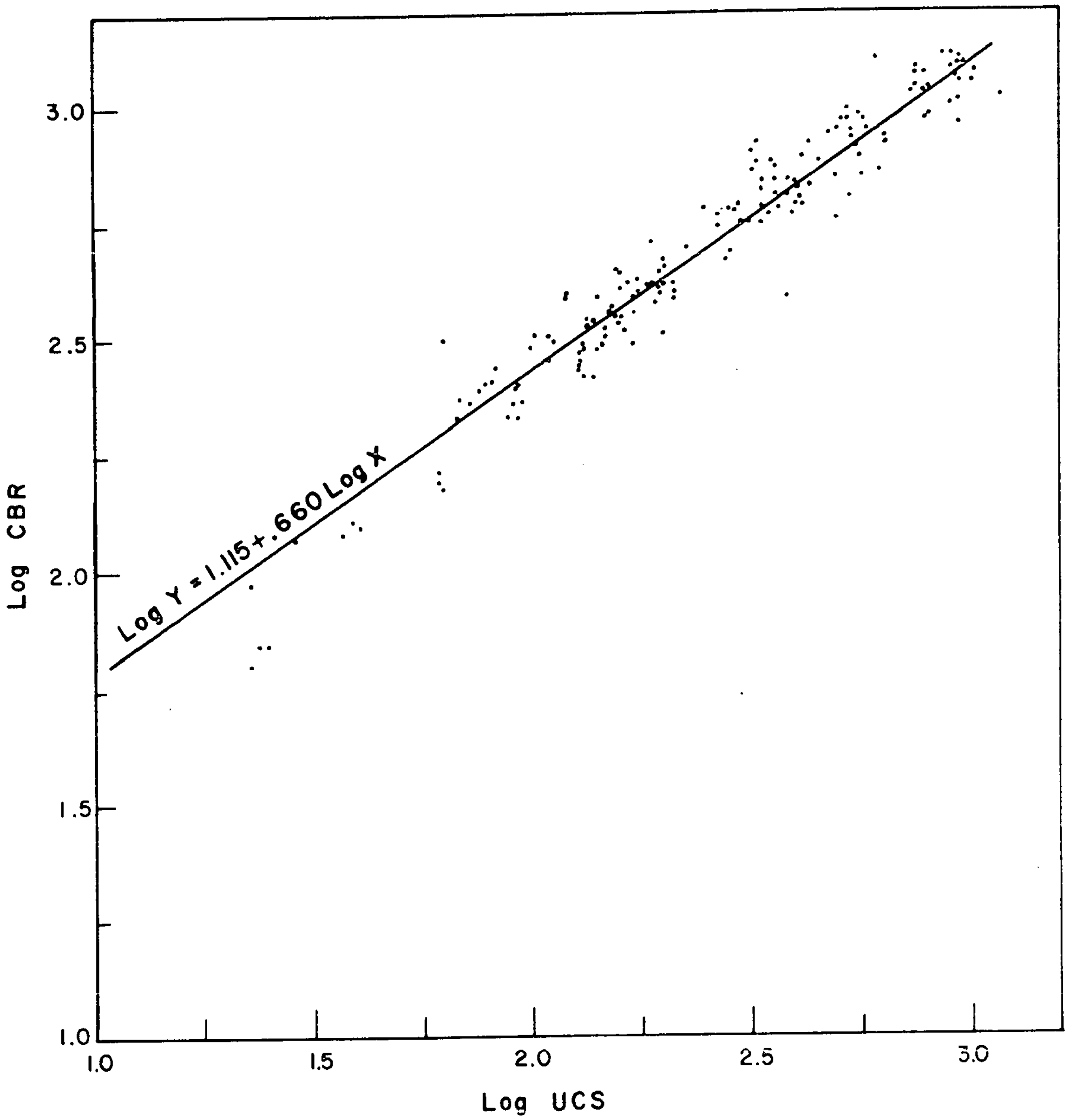

Log Y = 1.115 + .660 Log X
Log CBR
3.0
2.5
2.0
1.5
1.0
1.0
1.5
2.0
2.5
3.0
Log UCS

Figure 10. Relationship between the mean and the variance of UCS logs.

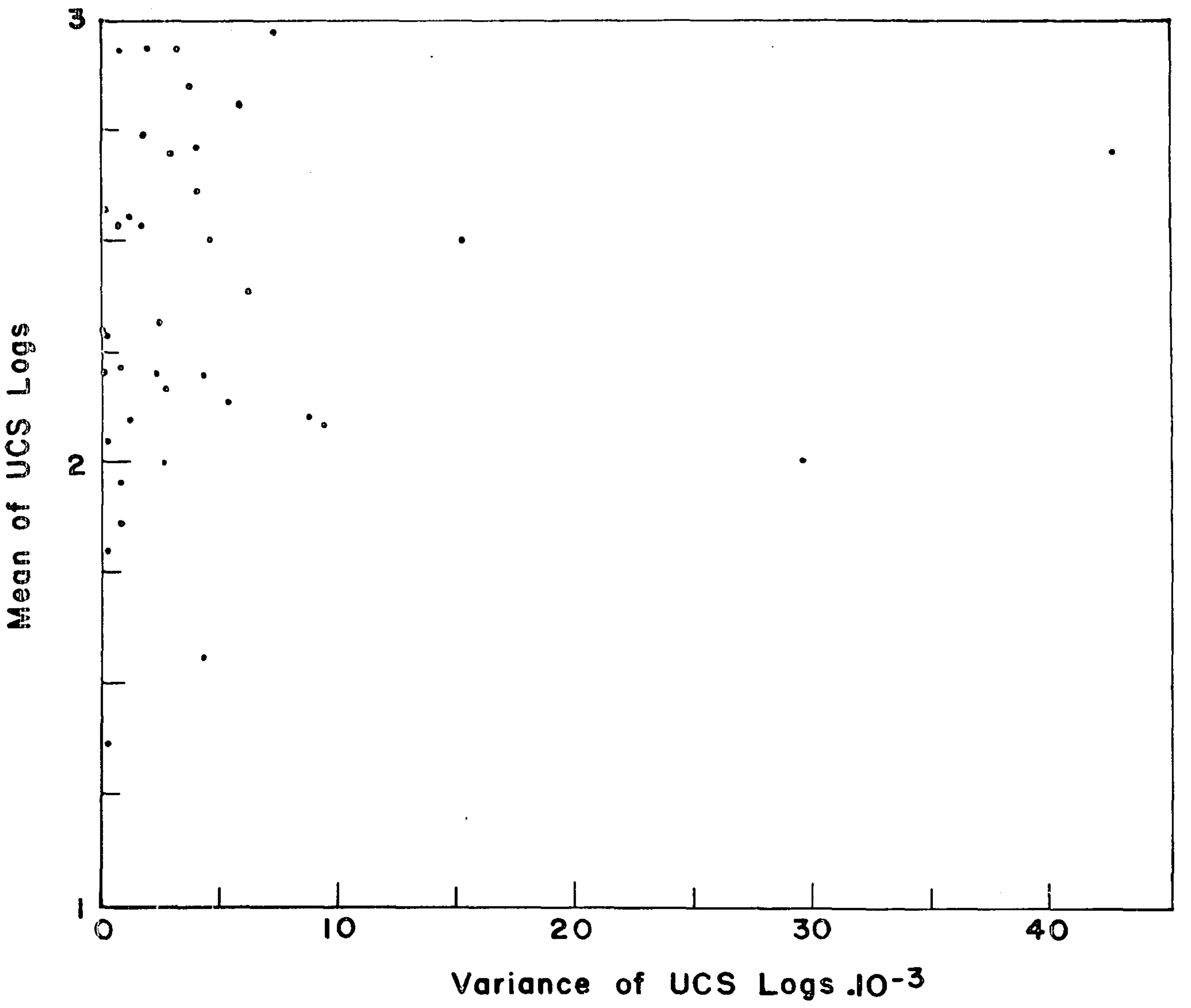

3
2
1
Mean of UCS Logs
0
10
20
30
40
Variance of UCS Logs .10^{-3}

Figure 11. Relationship between the mean and the variance of CRB logs.

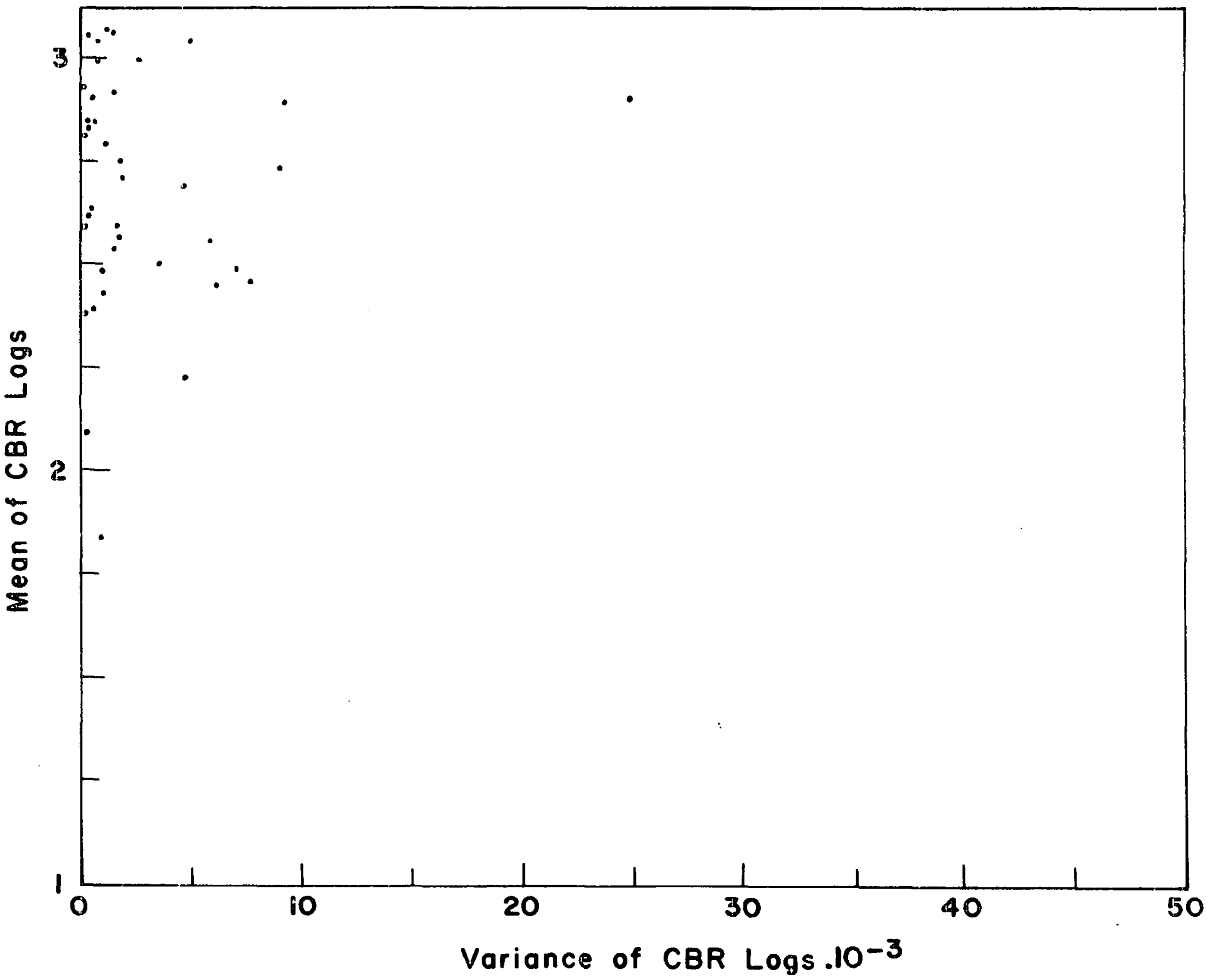

Mean of CBR Logs
3
2
1
0
10
20
30
40
50
Variance of CBR Logs .10^{-3}

Figure 12. Relationship between the mean of the UCS and CBR logs and the covariance of the UCS and CBR logs.

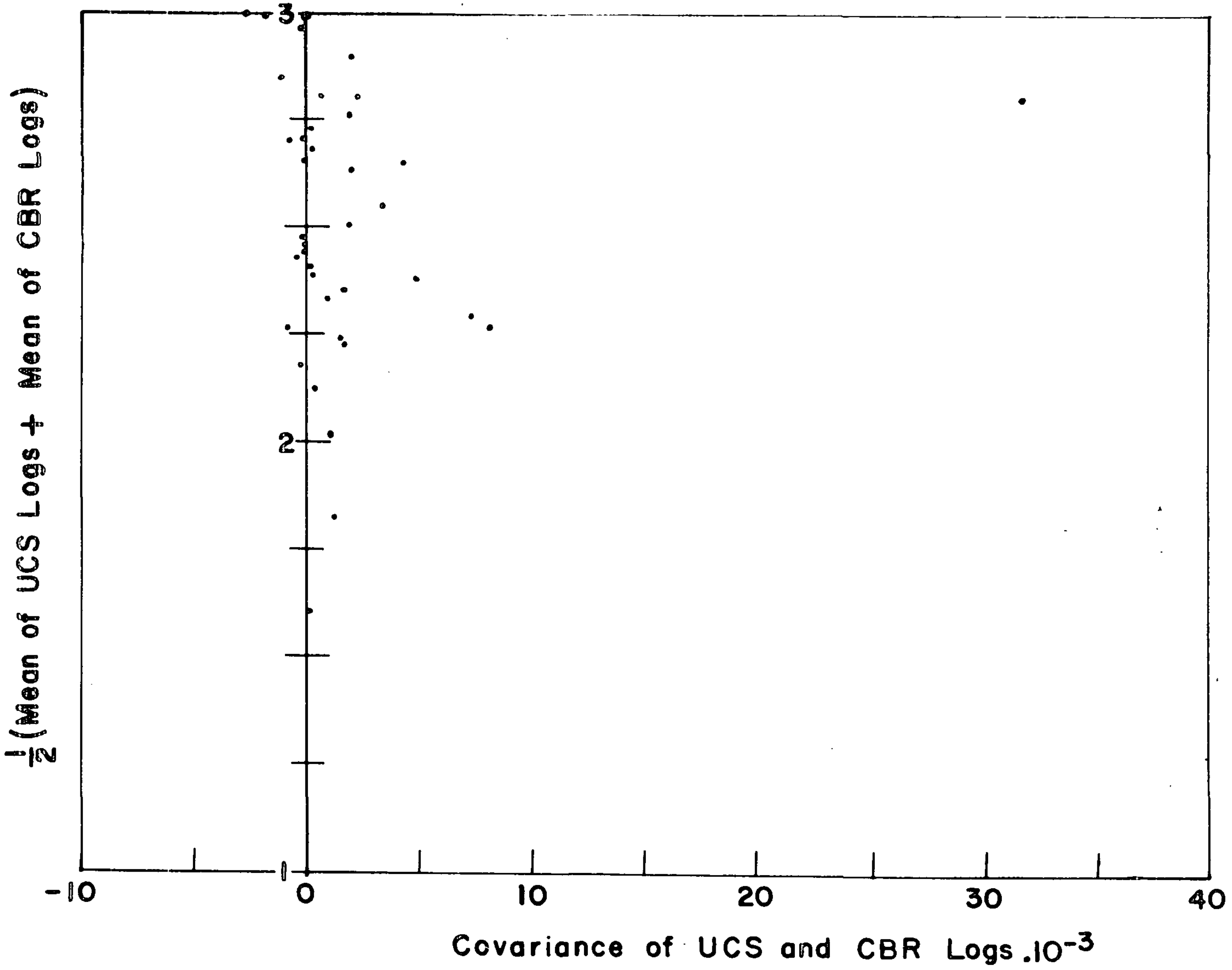
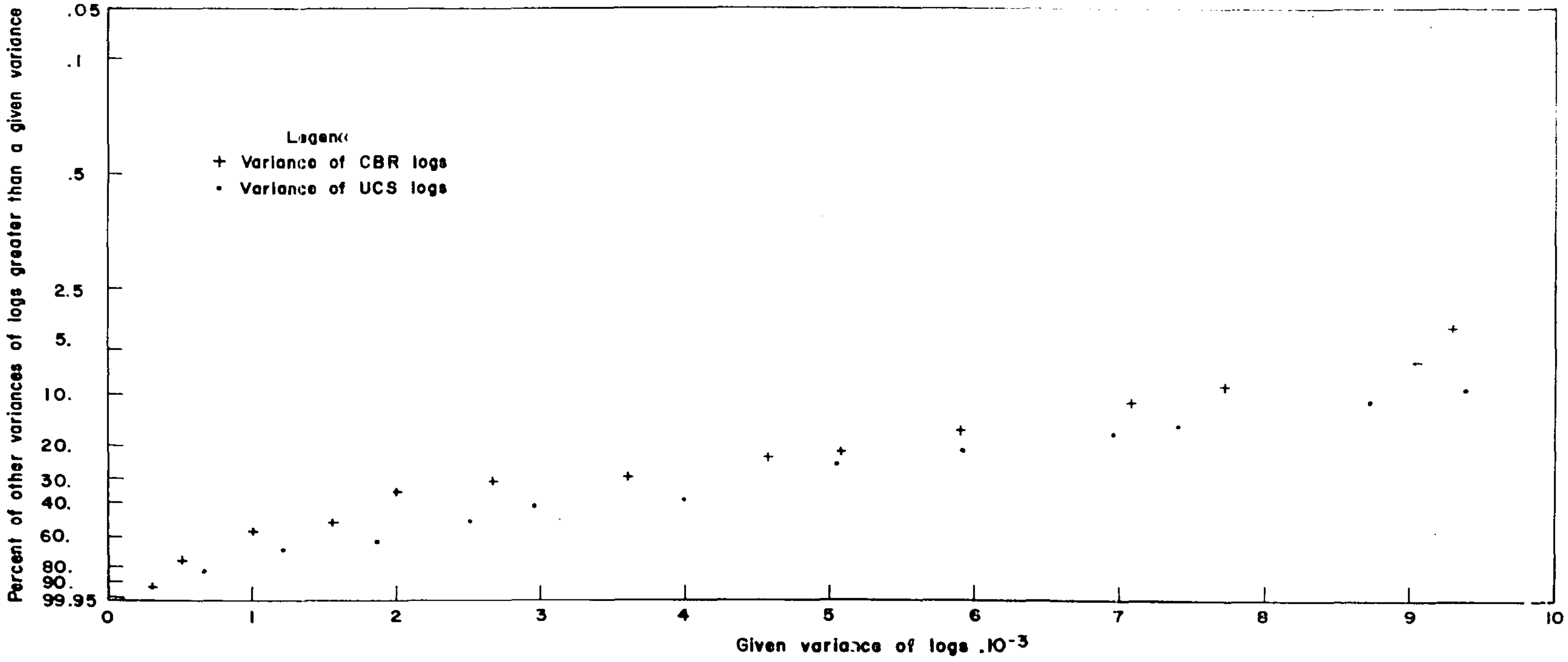

$\frac{1}{2}$(Mean of UCS Logs + Mean of CBR Logs)
3
2
1
-10
0
10
20
30
40
Covariance of UCS and CBR Logs .10^{-3}

Figure 13. Indication that the log CBR and log UCS variances follow the chi-square distribution arising under normality.


Percent of other variances of logs greater than a given variance
.05
.1
.5
2.5
5.
10.
20.
30.
40.
60.
80.
90.
99.95
Legend
+ Variance of CBR logs
• Variance of UCS logs
0
1
2
3
4
5
6
7
8
9
10
Given variance of logs .10^-3

$\alpha + \beta\xi_i$: errorless log CBR for the i th test condtion, a linear function of ξ_i

ϵ_{ij} : a normal error variable with mean zero and standard deviation σ_ϵ

η_{ij} : a normal error variable with mean zero and standard deviation σ

In addition, it is assumed in model (*) that the 312 error variables ϵ_{ij} and η_{ij} are uncorrelated except for a constant correlation between ϵ_{ij} and $\eta_{i'j'}$ when $i = i'$ and $j = j'$.

Figures 9, 10, 11, 12 and 13 suggest that the chosen model (*) is valid for the following reasons.

1. Figure 9 indicates that if a functional relation exists between "true" UCS logs and "true" CBR logs, then it is very likely to be a linear relationship.

2. Figures 10 and 11 show that it is not unreasonable to assume a constant variance for ϵ_{ij} and η_{ij}.

3. Figure 12 indicates that a constant correlation exists between ϵ_{ij} and η_{ij}.

4. Figure 13 suggests that the normality assumption is not unwarranted. It indicates that, to within expected sampling variation, the log CBR and log UCS variances follow the chi-square distribution arising under normality.

Assuming the model (*) to be acceptable, it is now possible to proceed with the construction of a 5% test of the hypothesis

$$H_o : \text{"}\alpha = \alpha_o, \beta = \beta_o\text{"} :$$

Compute

1. $\sum_i \sum_j (Y_{ij} - \alpha_o - \beta_o X_{ij})^2$

2. $\sum_j (X_{ij} - \bar{X}_{i.})^2/3 \equiv S^2_{X,i}$

$$\sum_j (Y_{ij} - \bar{Y}_{i.})^2 /3 \equiv S^2_{Y,\, i}$$

$$\sum_j (X_{ij} - \bar{X}_{i.})(Y_{ij} - \bar{Y}_{i}.) /3 \equiv S^2_{X,\, Y,\, i}$$

3. $S^2_X = \Sigma\, S^2_{X,\, i}/39$

$S^2_X = \Sigma S^2_{Y,\, i}/39$

$S^2_{X,\, Y} = \Sigma\, S^2_{X,\, Y,\, i}/39$

4. $$\mathcal{S}_{\alpha_o,\, \beta_o} \equiv \frac{\sum\limits_i \sum\limits_j (Y_{ij} - \alpha_o - \beta_o X_{ij})^2}{S^2_Y + \beta^2_o S^2_X - 2\beta_o S^2_{X,\, Y}}$$

Compare $\mathcal{S}_{\alpha_o,\, \beta_o}$ with $117 + [F_{39,\, 117}(.05)][39] \equiv K$

If $\mathcal{S}_{\alpha_o,\, \beta_o} < K$, accept H_o

If $\mathcal{S}_{\alpha_o,\, \beta_o} \geq K$, reject H_o

The above is a 5% test for the following reason.

Define

$$Z_{ij}(\alpha_o,\, \beta_o) \equiv (Y_{ij} - \alpha_o - \beta_o X_{ij})$$

Then

$$\mathcal{S}_{\alpha_o,\, \beta_o} = \frac{\sum\limits_i \sum\limits_j (Y_{ij} - \alpha_o - \beta_o X_{ij})^2}{S^2_Y + \beta^2_o S^2_X - 2\beta_o S^2_{X,\, Y}}$$

$$= \frac{\sum\limits_i \sum\limits_j Z^2_{ij}(\alpha_o,\, \beta_o)}{\dfrac{\sum\limits_i \sum\limits_j (Y_{ij} - \bar{Y}_{i.})^2 + \beta^2_o \sum\limits_i \sum\limits_j (X_{ij} - \bar{X}_{i.})^2 - 2\beta_o \sum\limits_i \sum\limits_j (X_{ij} - \bar{X}_i)(Y_{ij} - \bar{Y}_{i.})}{(3)(39)}}$$

$$= \frac{\sum_i \sum_j \left[Z_{ij}(\alpha_o,\beta_o) - \bar{Z}_{i.}(\alpha_o,\beta_o)\right]^2 + 4 \sum_i \bar{Z}_i^2 (\alpha_o,\beta_o)}{\frac{\sum_i \sum_j \left[Z_{ij}(\alpha_o,\beta_o) - \bar{Z}_{i.}(\alpha_o,\beta_o)\right]^2}{(3)(39)}}$$

$$= \quad (3)(39) \quad + \frac{4\sum_i \bar{Z}_{i.}^2 (\alpha_o,\beta_o)/39}{\frac{\sum_i \sum_j \left[Z_{ij}(\alpha_o,\beta_o) - Z_{i.}(\alpha_o,\beta_o)\right]^2}{(3)(39)}}$$

= 117 + (39) (a statistic distributed as $F_{39,\,117}$ under H_o). Q.E.D.

The above procedure is a "least squares" method for solving for the parameters of the model that is symetric in X and Y. In other words, this method will give the same answer whether Y is thought of as being regressed on X or X on Y. This is a natural requirement in the present situation since both X and Y are subject to error. This method is related to previous work on this type of problem which has been described in the literature (11, 12, 13).

Since the minimization of $\mathcal{S}_{\alpha,\beta}$ is much more difficult than the minimization involved in the usual least squares techniques, it was carried out by means of high speed computor. From the data thus obtained, it was found that $\min \mathcal{S}_{\alpha,\beta} = 2.73$. This value of 2.73 is an F-value well in excess of the 5% significance level. In fact, this value is very much in excess of the 1% significance level, thus leading to the assertion that the 99% confidence region for α, β is empty. This means that, at the 99% level, the model assumed for this data is not plausible.

The most suspect feature of the model (*) would seem to be the hypothesizing of a functional relationship between "true" UCS logs and "true" CBR logs. This then is a feature calling for re-examination.

Actually, it is somewhat more reasonable to think of a set of functional relationships, each corresponding to variation of but a single factor level, forming a two dimensional configuration in the plane bounded by two envelopes. It follows that, unless one is willing to specify factor conditions rather exactly, it becomes rather difficult to bring rigorous statistics to bear on the problem of determining a confidence interval for CBR values corresponding to specified UCS values.

The above does not, of course, preclude the possibility of using to good advantage the strong correlation evidenced in Figure 9 at least until further investigation yields statistical recipes as functions of factor conditions. It is to this end that the following equation of fit is presented:

$$\log Y = 1.115 + .660 \log X$$

where Y = CBR value and X = UCS value. The α and β values appearing in the equation, $\alpha = 1.115$ and $\beta = .660$, are in fact the α and β values that minimize $\mathcal{S}_{\alpha,\beta}$.

Engineering analysis of data

Figures 3 and 4 indicate the strength results obtained under varying conditions. As expected, strength values increased with increasing cement contents and increasing lengths of curing. In addition, as the length of curing increased, the rate of strength gain decreased. It is interesting to compare the immersed specimens with the unimmersed specimens. The immersed specimens had each an extra day of curing and as a result gained extra strength. On the other hand, the immersed specimens lost a certain amount of strength due to being immersed. For the 5% cement specimens the strength values are close to each other, indicating that the strength gained due to the extra day's curing is essentially nullified by the strength loss due to being immersed. With the higher cement contents, however, the immersion effect appears to be much more severe. It would seem as if this effect is mainly a function of length of curing. At low curing periods, immersion has little or no effect on strengths. At such times, the rate of strength gain is so fast that the extra day's curing tends to outweigh or balance the loss in strength due to being immersed. However as curing time increases, the rate of strength increase decreases and hence the strength loss due to being immersed is much more apparent.

From the statistical analysis of the data it appears that a true functional relationship, valid over a wide range of experimental conditions, does not exist between the unconfined compressive strength and the California Bearing Ratio. The equation, log CBR = 1. 115 + 0. 660 log UCS does, however, provide, a working relationship that can be used for rough predictions in investigations involving sand-cement mixtures. In addition, it is very possible that a true functional relationship may exist between the CBR and UCS for a given experimental condition where only one factor is varied at a time e. g. if the soil, cement content and method of curing are kept constant and only the length of curing is varied. In this sense, the above equation-although it cannot be considered to be an estimate of a single true relationship-can be considered to be the "average" of many single factor relationships. If this be true, then it could be further hypothesized that a true relationship may exist between the CBR and the UCS of stabilized soils where the only variable is the soil type. In regarding this hypothesis, however, it should be kept in mind that soil type is not as well defined a factor as either cement content, curing time or method of curing.

A word deserves to be said regarding a strength criterion for sand-cement. A commonly acception criterion is that of an immersed strength of 250 psi after 7 days moist curing (3). Based on the results obtained in this investigation, it appears that a sand-cement mixture with an unconfined compressive strength of 250 psi has a California Bearing Ratio of about 500 percent. Similarly, a sand-cement mixture with a CBR of 120 percent has a UCS value of about 29 psi. These figures immediately suggest that a criterion of 250 psi for stabilized sand is unreasonable as it fails to take into account the inherent strength due to lateral confinement.

Part 2 - Detecting Outliers in a Large Series of Soil-Additive Strength Determinations

As mentioned before, he unconfined compressive strength test is probably the most commonly used test in soil stabilization investigations. The general procedure is, for one given test condition, to prepare and test several specimens, after which the average of the several strength values is reported. Three specimens per test condition are commonly used. Because of the many variables involved, the total number of specimens which may have to be tested may range from the hundreds to the thousands, depending upon the size and scope of the investigation.

Since such large numbers of specimens are involved, it is likely that some unconfined compressive strength results will be obtained that are, seemingly, not what they should be. The question then arises whether these unusual observations are the result of expected normal experimental variation, or whether they are due to an experimental or material aberration and should therefore be discarded. In cases where three specimens are prepared per test condition, a commonly used solution to this question is to discard any single measurement which deviates by more than ten percent from the average of all three measurements, as prescribed in ASTM "Method of Test for Compressive Strength of Hydraulic Cement Mortars" (2). In the event of such a disqualifying deviation, the average of the remaining two strength values is then reported.

It is felt that this blanket-type disqualifying percentage should be reappraised from a statistical point of view, since it is very possible that entirely valid triplicate unconfined compressive strength values may attain this percentage simply by virtue of expected statistical fluctuation. Thus many values may be unjustly disqualified. Since unjustly disqualified strength values carry information which is as valid as that carried by their supposedly more reliable neighbors, uncritical adherence to such a blanket-type disqualifying percentage causes needless loss of information. In addition, bias is introduced when any strength observation is wrongfully discarded.

Purpose of the study

In the previous experiment, there was detailed a test whereby outliers in a series of quadruple strength determinations-where each member of the quadruple came from a different batch-could be detected. The purpose of this phase of the investigation, therefore, was to present a procedure by which outliers in a series of triplicate strength determinations-where each member of the triplicate came from the same batch-could be detected. In addition, it is often desired to have some criterion by which to judge an investigation as a whole and therefore a method is given for examining a series as a whole for reliability, homogeneity and normality.

Disqualification test for triplicate studies

The statistical theory of the present approach requires the existence and the estimation of a constant coefficient of variation-abbreviated CV-for the entire series of observations. The CV of any observation equals the dispersion to which that observation is subject divided by the true value that the observation is supposed to estimate. It should be a constant for all the observations of a single investigation.

A simple nomographic procedure has been devised for establishing and estimating this constant CV.

Procedure for estimating the CV:

1a. For each set of triplicate unconfined compressive strength values, compute the ratio, r, of the range, R, of the three values to the average, $\bar{X}$, of the three values. The range is defined as the difference between the largest value and the smallest value of the three. Thus

$$r = \frac{R}{\bar{X}} = \frac{X_{max} - X_{min}}{(X_1 + X_2 + X_3)/3}$$

1b. Arrange all the r values so obtained in ascending order of magnitude. This can easily be done by plotting them on ordinary graph paper.

1c. Choose approximately thirty well spaced r values. For each selected r value, find the number, n, of other r values less than it, add

1/2 to this,[a] and express this number as a percentage of the total number, N, of r values; that is, compute

$$100(n + 1/2) /N$$

1d. Plot each percentage against its corresponding r values on the nomograph in Figure 14, using scale A for the r values and scale B for the percentages.

1e. Fit the thirty points so obtained with a straight line-hereafter called the CV line-passing through the origin. If the points lie reasonably close to the straight line, then constancy of the CV is established and the proposed test is applicable. (Questions of objective fit and closeness criteria are touched upon in the discussion).

Outliers, if present, will tend to unduly enlarge r. This will cause the r pattern to form an arched rather than straight line. In such cases, the points furthest from the origin should be excluded from the straight line fit. A technical though perhaps impractical refinement here is to eliminate far points until the remaining replotted points form a satisfactory straight line.

The CV itself is estimated by the value on scale A at which the CV line attains a height of 24 on scale B.

It might be noted that prior workers in this general area have worked with the assumption of a constant CV (14). In addition, a considerable number of experimental sets of data have been examined for constancy of the CV at the Iowa Engineering Experiment Station, and it has been found to hold in every case.

Upon the establishment and estimation of the constant CV, it is now possible to test for possible incorrect unconfined compressive strength values. The procedure is as follows:

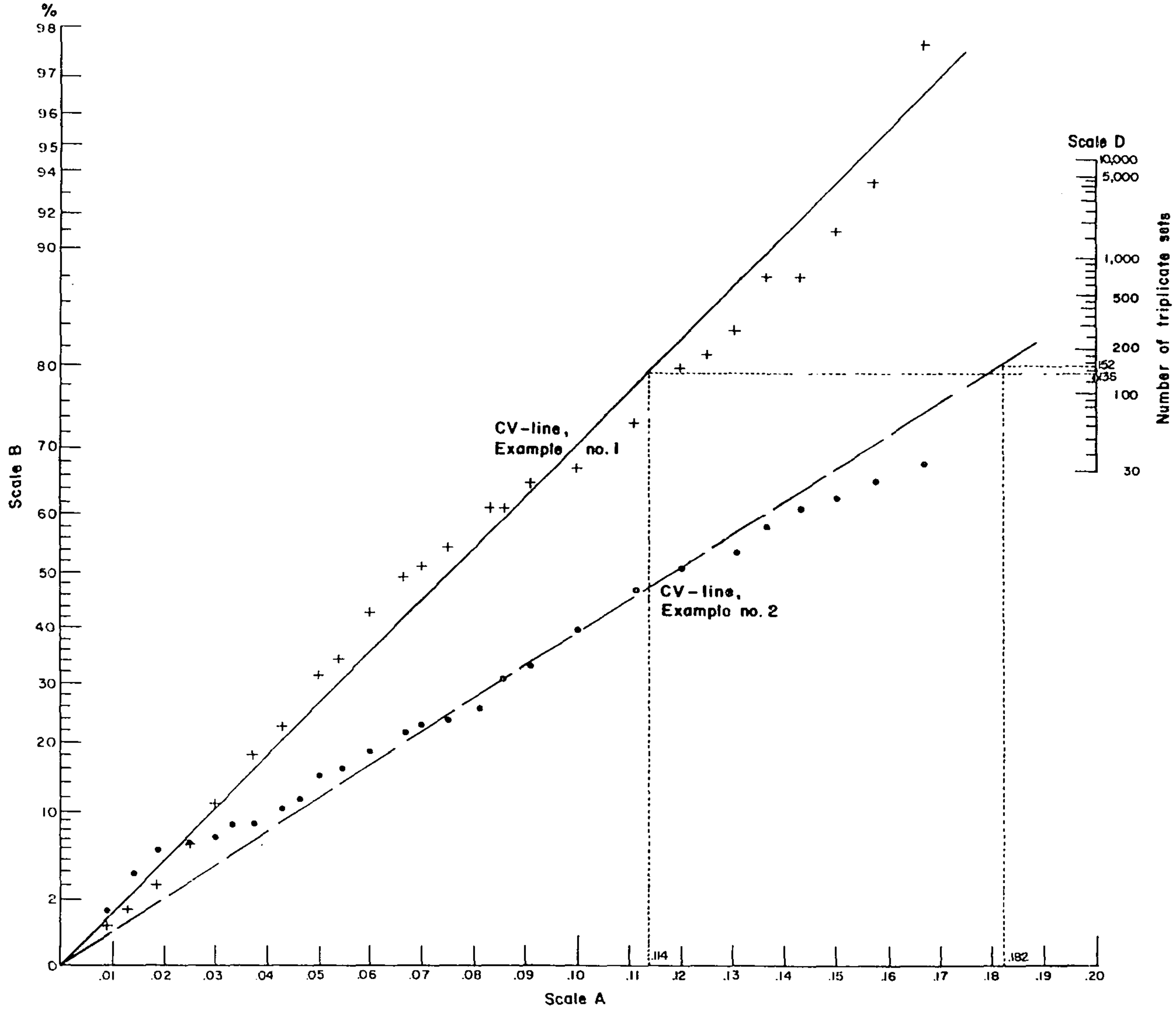
Procedure for disqualification of extreme strength values:

2a. For each set of triplicate values compute the ratio U of the largest value (X_{max}) - the average value ($\overline{X}$) to the average value ($\overline{X}$).

---

[a] For N values greater than 100, it is not necessary to add 1/2.

Figure 14. Nomographic computation of disqualifying critical values for triplicates of specimens.


%
98
97
96
95
94
92
90
80
70
60
50
40
30
20
10
2
0
Scale B
CV-line, Example no. 1
CV-line, Example no. 2
Scale D
10,000
5,000
1,000
500
200
152
135
100
30
Number of triplicate sets
.01
.02
.03
.04
.05
.06
.07
.08
.09
.10
.11
.114
.12
.13
.14
.15
.16
.17
.18
.182
.19
.20
Scale A

Thus

$$U = \frac{X_{max} - \bar{X}}{\bar{X}}$$

2b. For each set of triplicate values, compute the ratio V of the average value ($\bar{X}$) - smallest value (X_{min}) to the average value ($\bar{X}$). Thus

$$V = \frac{\bar{X} - X_{min}}{\bar{X}}$$

2c. Enter scale D in Figure 14 at the total number of triplicate sets. Through this point draw a horizontal line until it intersects the CV line through the origin. Read on scale A the value t of the abscissa of this intersection point.

2d. t is the critical value for both U and V. Any triplicate whose U exceeds t should have its X_{max} discarded; similarly, any triplicate whose V exceeds t should have its X_{min} discarded. In other words the t value, when expressed in percentage form, is the disqualifying percentage for the investigation at hand.

It must be realized that, although the suggested procedure controls the rate of wrongful disqualifications, it cannot reduce this rate to zero. It is therefore possible that valid observations may be disqualified. Similarly, a certain number of outliers will not be detected. Wrongful disqualifications can occur either when all three members of the triplicate set are subject only to normal experimental variation or possibly because the two remaining values are, in fact, the illegitimate ones. The investigator seeking additional controls for errors of this type may wish to cross check the disqualifications suggested by the present procedure against the disqualifications suggested by the magnitude of the corresponding residuals from fitted regression functions (15). This cross check is a standard statistical test and is not further discussed in this study.

Where, however, the cross check is not used, it is recommended that, if one observation is disqualified, the middle observation of the original three then be reported. If it should happen that both U and V are extreme for one triplicate set, the entire triplicate set should then be discarded.

Criterion for the reliability of the investigation as a whole

In some cases it may be of interest to check on the reliability of the investigation as a whole. This may be necessary for many reasons, such as suspected unreliability of the operator, non-normality, or inhomogeneity of the material under test.

Proposed reliability test:

3a. Arrange all the U values in ascending order of magnitude. This is most easily done by plotting them on ordinary graph paper.

3b. Select approximately thirty well spaced U values. For each selected U value, find the number of other U values that are less than the selected U value and express this number as a percentage of the total number of U values.

3c. Using the nomograph in Figure 15, plot on scale E each percentage obtained in 3b against its corresponding U value on scale A.

3d. Fit the points so obtained by a straight line-hereafter called the U line-through the origin.

3e. Similarly, do 3a, 3b, 3c, and 3d for V so as to obtain a V line.

The extent of non-coincidence of the three lines obtained in 1d, 3d, and 3e, and the extent to which the three sets of points fail to be fitted by the CV line, indeed the actual shape of the sets themselves, will provide clues concerning series-wide unreliability, inhomogeneity and non-normality. For example, inhomogeneity, in the sense of more than one underlying coefficient of variation, will cause the three sets to form similar "S" shaped curves, arching first downward then upward, the first arch typically being the more pronounced. This effect is similar to that arising under "inadvertent plot splitting" in half-normal plot analyses (16), and is due to similar causes. Again, certain types of operator fabrication will manifest themselves in distinctive patterns. For example, fabricating a triplicate from a single determination by adding and subtracting fixed proportions of the single determination will cause a vertical discontinuity to appear in all three plots. On the other hand, fabricating a triplicate from a pair of determinations by interpolation will cause a configuration similar to but typically less extreme than that arising under inhomogeneity.

Should serious series-wide non-normality be uncovered, the clash of non-normal data with normal theory should, as a rule, be resolvable in favor of the theory. In other words, non-normality of data often will have an identifiable and removable cause.

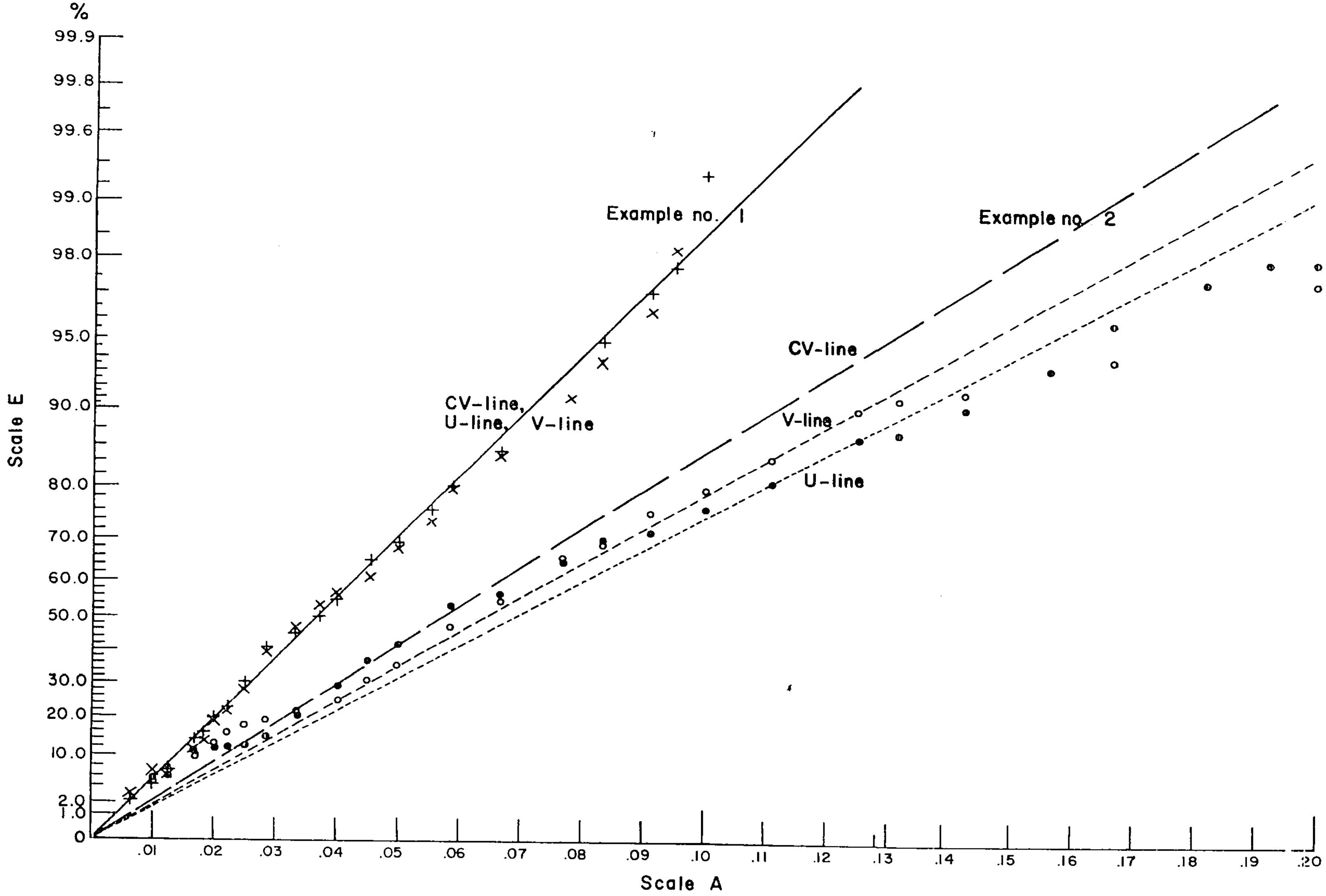
Examples involving the use of the recommended procedures

To illustrate the use of the proposed techniques, data from two typical studies taken from the files of the Iowa State University Engineering Experiment Station were analyzed. The first example involved 134 triplicate sets of unconfined compressive strength determinations of soil-calcium lignosulfonate-aluminum sulphate specimens (17). The second example involved 152 triplicate sets of unconfined compressive strength determinations of soil-lime-sodium silicate specimens (18). The recommended procedures were applied to this data as indicated in Figures 14 and 15. The values and calculations upon which the example no. 2 graph is based are shown in Reference 18.

As shown in Figure 14, the estimated CV for the first example is 0.048, and the critical t is 0.114, corresponding to a disqualifying percentage of 11.4. None of the 134 triplets were disqualified by this criterion. As shown in Figure 15, the CV line and V line coincide, with the U points and V points falling close to this joint line. All indications therefore point to the fact that this investigator was in thorough control of his experiment.

The estimated CV for the second example is approximately 0.074, indicating a degree of experimental precision lower than that of the first example. This lower precision probably does not represent an operator effect, but is probably due to the well known rapid jell-forming

Figure 15. Nomographic assessment of series reliability.


%
99.9
99.8
99.6
99.0
98.0
95.0
90.0
80.0
70.0
60.0
50.0
30.0
20.0
10.0
2.0
1.0
0
Scale E
.01
.02
.03
.04
.05
.06
.07
.08
.09
.10
.11
.12
.13
.14
.15
.16
.17
.18
.19
.20
Scale A
Example no. 1
CV-line,
U-line, V-line
Example no. 2
CV-line
V-line
U-line

ability of sodium silicate. Low precision does not by itself constitute evidence of experimental inefficiency but, as is likely in the present case, can be the result of inherent material properties. The critical t-value for this example is approximately 0.182, corresponding to a disqualifying percentage of 18.2. As regards the reliability check carried out in Figure 15, the CV line, U line and V line are seen not to coincide. Moreover, the U points and V points do not lie close to their respective lines. The tendency to downward curvature exhibited by both the U points and V points suggests the possibility of inhomogeneity of experimental material.

It is important to note that the critical percentage of 11.4 for the first experimental series is near the commonly accepted blanket percentage of 10%, which, parenthetically, is exceeded by 3 triplicate sets of this series. This 10% is also exceeded by 38 triplicates of the second series. Use of the critical percentage "tailor-made" to inherent experimental variability thus leads to a reduction in the number of disqualification in the case of both experimental series. These are, namely, zero versus 3 for example No. 1 and 18 versus 38 for example No. 2.

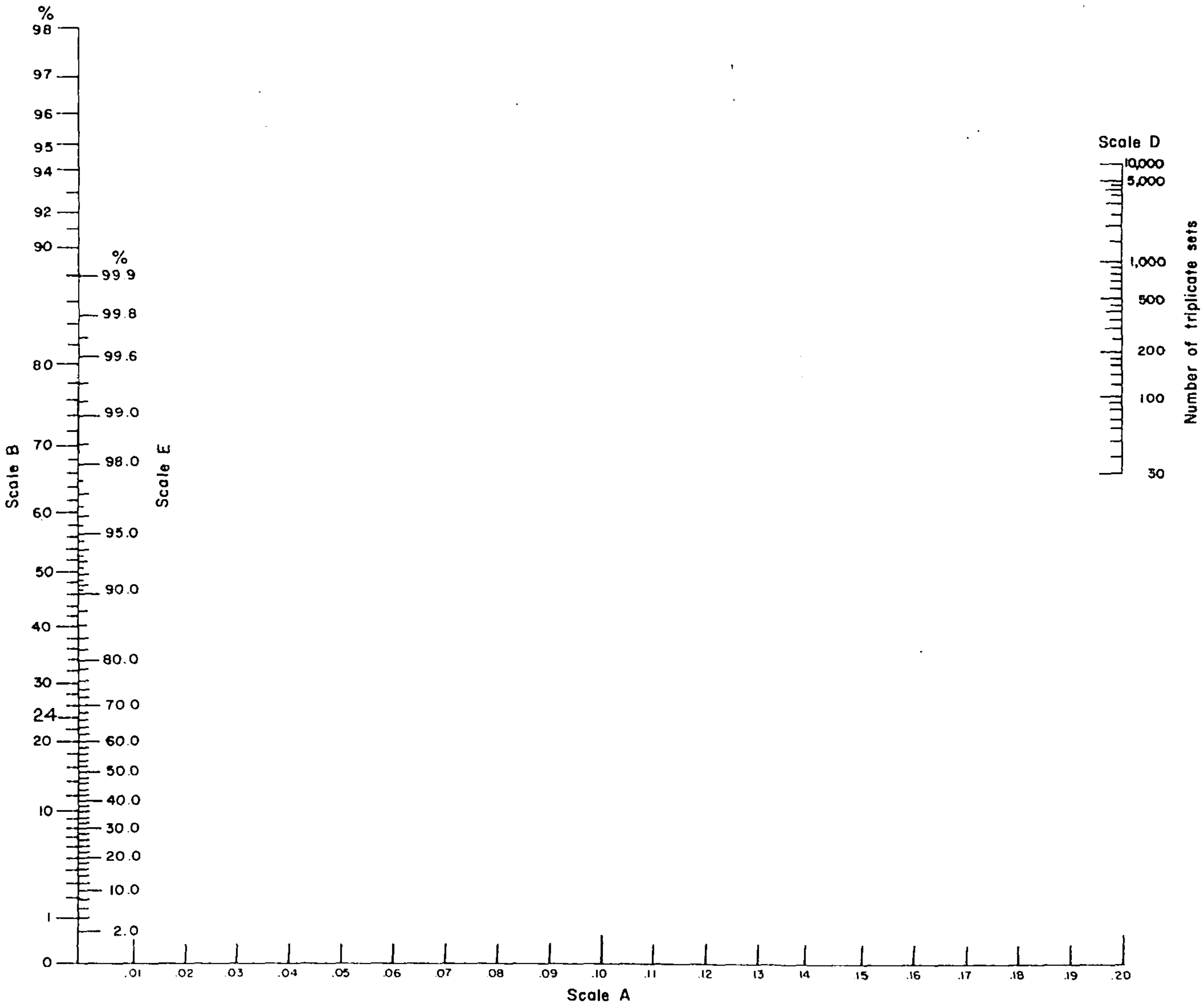
Note that the two types of nomographic computations shown in Figures 14 and 15 can be performed on a single nomograph. A sample of such a nomograph, called "Outlier Paper" is given in Figure 16.

Discussion

The Outlier Paper of Figure 16 is based upon the following facts.

A. The ratio $\frac{r}{CV}$ has approximately the distribution of the range of three unit normal deviates (19), and $\frac{U}{CV}$ and $\frac{V}{CV}$ have approximately the distribution of the largest minus the average of three unit normal deviates (20). Verifying computations indicate that these approximations are sufficiently exact as long as the coefficient of variation is less than 0.15. Scales A and B represent inverse probability transformations corresponding to the above two functions of unit normal variables (19, 20). The linearizing property of inverse probability transformations has been exploited before (16).

Figure 16. Outlier paper for triplicates of specimens.


%
98
97
96
95
94
92
90
80
70
60
50
40
30
24
20
10
1
0
Scale B
%
99.9
99.8
99.6
99.0
98.0
95.0
90.0
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
2.0
Scale E
.01
.02
.03
.04
.05
.06
.07
.08
.09
.10
.11
.12
.13
.14
.15
.16
.17
.18
.19
.20
Scale A
Scale D
10,000
5,000
1,000
500
200
100
30
Number of triplicate sets

B. In view of the above, the cumulative distribution functions for r, U and V are straight lines through the origin and have slope of $\frac{1}{CV}$ when plotted on the outlier paper. This enables the CV line, which is in fact the estimated cumulative distribution of r, to yield critical values for U and V, i. e. to be used as if it were in fact the cumulative distribution function of U and V.

It is important to note that, ideally, the construction of the CV line should be based on a statistic that is as insensitive as possible to outliers, whereas the disqualifying percentage derived from this CV line should be applied to statistics that are as sensitive as possible to outliers. Triplicate observations may lend themselves only partially to these objectives if, as is assumed in this paper, both large and small outliers are involved. In view of this, the plot of the partially sensitive r values may show some downward curvature. In such cases, as has already been recommended, the CV line should be fitted on the basis of the r points less likely to be contaminated by the outliers, i. e. the r points closer to the origin.

In cases where it is known that only large outliers are present, an ideal insensitive statistic is the ratio of the difference to the mean of the middle and smallest observation.

C. The method of obtaining the disqualifying percentage is based upon the "multiple-comparison" point of view that experimental series not containing outliers, regardless of their length, should suffer no disqualification with probability 0. 5. It is realized that other points of view regarding the question of risk will lead to different D scales.

It is of interest to note the manner in which the critical disqualifying values for U and V depend upon the total number of triplicate sets and also upon the constant coefficient of variation. When the number of triplicate sets increases, the critical t value increases, which means that the critical U and V values also increase. This follows from the present point of view regarding risk and may be explained by the fact that, since a greater number of triplicates are involved, natural experimental variation is expected to produce greater numbers of extreme U and V values. The critical t value also increases with increasing CV.

This is a reflection of the fact that the data are expected to be more erratic whenever the natural experimental error, of which the constant CV is a measure, is large.

Further theoretical considerations revolve about the manner of fitting the CV line and the manner of assessing the goodness-of-fit of the r, U and V points to this line. As a rule, an eye-fit will be adequate for the CV line, as other more sophisticated methods probably will not provide sufficiently greater accuracy to compensate for their greater computational complexities. A measure of goodness-of-fit is provided by the maximum vertical deviation, in units of percentage, of the thirty points from the straight line. This deviation may be approximately judged in terms of the known distribution of the maximum vertical discrepancy between a population CDF and its corresponding sample CDF (10). However this distribution theory should be taken only as a rough guide since (a) only thirty points of the sample CDF have been plotted, (b) the CDF to which this sample CDF is being compared is a fitted rather than a true CDF and (c) whatever outliers are present are actually contributing to the discrepancy between the two CDF's; alternatively, if one attempts to eliminate outliers by the refinement given in 1e, maximum vertical deviations will arise that are considerably smaller than those expected according to the standard distribution theory.

Part 3 - Further Methods for the Control of Data Quality

Many are the problems confronting the soil engineering investigator who is about to begin a study. One problem always before him is how many specimens should he prepare per test condition. Some investigators like to use four specimens and this, of course, poses the problem of how to detect outliers in the results obtained from their studies. Since large experiments-such as the cement-fly ash one described in Reference (21)-are carried out over a long period of time, there is always the to-be-feared possibility that certain time-associated biases, perhaps due to operator or apparatus deterioration, may creep into the work and thus taint the results.

These are but a few of the problems and decisions confronting the investigator. It cannot be expected of him that he can solve all of them correctly-it can only be hoped that he can know and minimize his errors.

Purpose of the study

The purpose of this investigative phase was to develop procedures that would be helpful to the soil engineer in overcoming some of these problems. The following is a brief listing of the items discussed.

1. Outlier test for studies involving four specimens per test condition.
2. The advantages-cum-disadvantages of using four, instead of three, specimens per test condition.
3. Method of selecting specimens so as to minimize time-affected biases.
4. The use of control specimens in evaluating data quality.
5. The influence of inadequate preparation of the soil sample.

Proposed disqualification test for studies involving four specimens per test condition

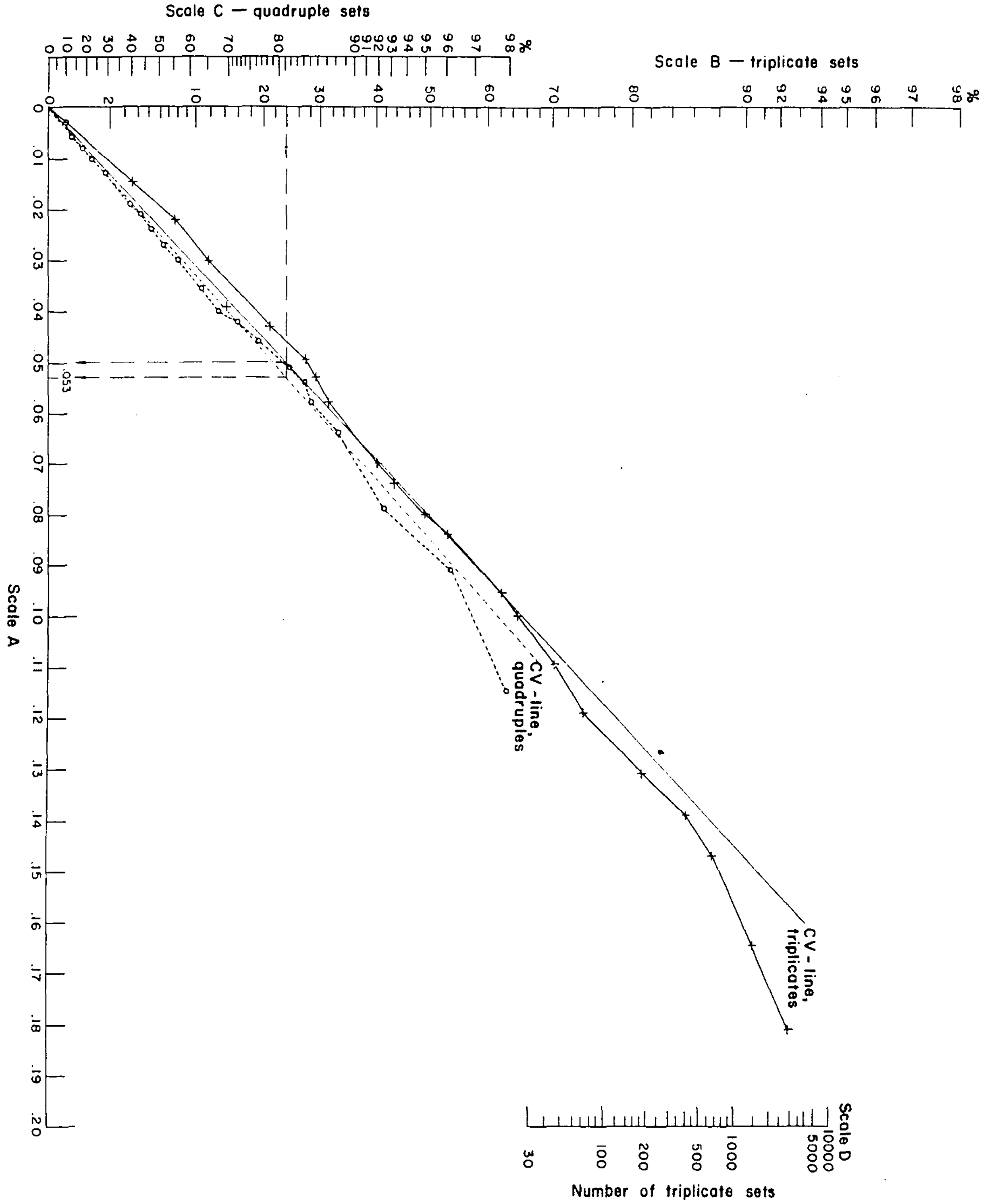
As indicated in Part 2, the statistical theory of this approach requires the existence and the estimation of a constant coefficient of variation (CV) for the entire series of observations. The CV of any observation equals the dispersion to which the observation is subject divided by the true value that the observation is supposed to estimate. This should be a constant for all the observations of a single investigation. The following is a simple nomographic procedure for establishing and estimating this constant coefficient of variation.

Procedure for establishing and estimating the CV:

1a. For each set of quadruple unconfined compressive strength values, compute the ratio, r, of the range, R, of the middle two strength values to the average $\bar{X}$, of these two same values. Thus,

$$r = \frac{R}{\bar{X}} = \frac{X_2 - X_3}{(X_2 + X_3)/2}$$

Figure 17. Nomographic computation of disqualifying critical values for quadruples of specimens.


Scale C — quadruple sets
0 10 20 30 40 50 60 70 80 90 91 92 93 94 95 96 97 98 %
Scale B — triplicate sets
0 2 10 20 30 40 50 60 70 80 90 92 94 95 96 97 98 %
Scale A
0 .01 .02 .03 .04 .05 .06 .07 .08 .09 .10 .11 .12 .13 .14 .15 .16 .17 .18 .19 .20
.053
CV - line, quadruples
CV - line, triplicates
Scale D
30 100 200 500 1000 5000 10000
Number of triplicate sets

1b. Arrange all the r-values so obtained in ascending order of magnitude. This is most easily done by plotting them on ordinary graph paper.

1c. Choose approximately 30 well spaced r values. For each selected r value, find the number, n, of other r values less than it, add 1/2 to this[a], and express the resulting figure as a percentage of the total number, N, of r values-that is, compute

$$100(n + 1/2) /N$$

1d. Plot each percentage thus obtained against its corresponding r value on the nomograph given in Figure 17. Use scale A for the r values and scale C for the percentages.

1e. Fit the points so obtained with a straight line-hereafter called the CV line-which passes through the origin. If the points lie reasonably close to the assumed straight line, then constancy of the CV is established and the proposed test is applicable.

Obviously the construction of the CV line should be based on a statistic that is as insensitive as possible to outliers. With quadruple observations, almost complete insensitivity is achieved by using the above mentioned ratio of the difference to the average of the middle two observations. However some outliers may be present which may tend to enlarge r unduly. This situation will cause the r pattern to form an arched rather than a straight line. In such cases, the points which emphasize this arch i. e. the points farthest from the origin, should be excluded from the straight line fit.

The CV itself is estimated by reading the value on scale A at which the quadruple CV line attains a height of 24 on scale B.

Upon the establishment and estimation of the constant CV, it is now possible to check for possibly incorrect strength values.

Procedure for disqualification of extreme strength values:

2a. For each set of quadruple values, compute the ratio U of the largest value, X_{max}, minus the average of the middle two values to the

[a] For N values greater than 100, it is not necessary to add 1/2.

average of the two middle values. Thus

$$U = \frac{X_{max} - \bar{X}}{\bar{X}}$$

2b. For each set of quadruple values, compute the ratio, V, of the average value of the middle two values, $\bar{X}$, minus the smallest value, X_{min}, to the average of the middle two values, $\bar{X}$. Thus

$$V = \frac{\bar{X} - X_{min}}{\bar{X}}$$

2c. Enter scale E at the total number of quadruple sets. Through this point draw a horizontal line until it intersects the CV line through the origin. Read on scale A the value, t, of the abscissa of the intersection point.

2d. For both U and V, this t value is the critical value. Any quadruple whose U exceeds t should have its X_{max} discarded; similarly any quadruple whose V exceeds t should have its X_{min} discarded. In other words, the t value-when expressed as a percentage-is the disqualifying percentage for the investigation at hand.

It is recommended that if the U and/or V value(s) are suspected of being extreme by this procedure, that the average of the middle two observations be then reported. If neither the U or V value is suspected, then the average of all four values of a set should be reported as being the "true" value.

This method of obtaining the disqualifying percentage is based upon the "multiple comparison" point of view that experimental series not containing outliers, regardless of their length, should suffer no disqualification with probability 0.5. In other words, if the series is entirely clean and no outliers are present, there is a 50/50 chance that no data will be disqualified.

The disqualifying percentage, i.e. the critical disqualifying values for U and V, depends upon the total number of quadruple sets and also upon the constant coefficient of variation. When the number of quadruple sets increases, the critical disqualifying percentage-that is the t value-also increases. This seemingly odd fact can be explained by the fact that since a greater number of quadruple sets are involved, greater numbers of extreme but valid U and V values can be expected because of natural experimental variation. In addition, the critical t value also increases as the CV increases. This can be explained by the fact that the data are expected to be more erratic wherever the natural experimental error, of which the constant CV is a measure, is large.

Discussion This section describes a method of detecting outliers in series of strength determinations involving quadruples. In Part 1 there was also described another method of detecting outliers in a series involving quadruples. The question then naturally arises as to why two methods are presented and what are the relative merits of each.

To clearly understand this, it is necessary to look at both methods of obtaining data. In the UCS-CBR work, members of a quadruple were obtained-at the same factor level-from different batches. As a result, it is necessary to have a technique to detect outliers that makes no assumption as to how the variance varies on the basis of μ . This is what was done in this case and as a result the proposed method has the added effect of being able to detect any "split-plotting" as expressed by $(\mu - \mu \text{———} \mu - \mu)$.

The above is not true for the problem discussed in this section, where all four specimens came from one batch. Since this method is not expected, and is not able, to detect any batch to batch difference-but rather is a measure of any man or machine measurement error during the course of the experiment-the assumption is made that the variance is a known function of μ times an unknown constant. This can be written in the form

$$\sigma^2 = K \mu^2$$

This unknown constant is the $(CV)^2$ when the known function is μ^2 . It is the variance if the known function is 1. This assumption does not include the situation where the variance is made up of two components-one of which is subject to a constant CV and the other to a constant variance-as illustrated in the UCS-CBR method. In that problem, the members of each quadruple, although independent, were subject to both a within and a between-batch component. In that case, the variance can be expressed in the form

$$\sigma^2 = A + B\,\mu^2$$

where both A and B are unknown. This obviously is not of the form

$$\sigma^2 = K\,\mu^2$$

as illustrated in this section.

As of now, a technique has not been developed which is suitable for both types of problems. Hence it is necessary to present both methods of analyses.

Observations on the number of specimens per test condition

A problem always confronting the research engineer is how many specimens should he use per test condition. Obviously, the more he uses, the more confident he is of his data and conclusions. However, practical economics dictate that he keep the numbers as low as possible. In large soil engineering studies involving the unconfined compressive strength test, it is common practice to prepare at least three specimens per test condition. One could argue that if three specimens give good results, then four would be better and it would not be too much trouble to prepare an extra specimen per test condition. Five or more specimens would, of course, be even better still, but use of these numbers could increase the burden of work in a large investigation by tremendous amounts. Hence it was decided to use four samples as a comparison with three samples in order to determine if the extra sample increased precision by a worthwhile amount. Since the CV is an indirect indication of precision, it is used as a basis for comparing the two sets of values.

To provide an illustrative example for this study, an experiment was devised which involved two soils, three percentages of a portland cement, and four percentages of each of three different fly ashes. The non-statistical analysis of this experiment is described in reference (21). For each combination of soil, cement and fly ash, four specimens were prepared and compacted at their optimum moisture content for maximum density. In order to determine the CV for the series, use was made of the nomograph described in Figure 17 and the procedure as detailed for quadruples. The CV line that was obtained is shown as the dotted line in Figure 17. The strength values and calculations upon which this line is based are shown in Appendix B. A method has already been described in Part 2 by which outliers can be detected in a series of soil-additive strength determinations involving three strength values per test condition. The suggested procedure is very similar to that described here for quadruplets, except that the statistic used in establishing the triplicate CV is

$$\frac{X_1 - X_3}{(X_1 + X_2 + X_3)/3}$$

and that for the quadruple CV is

$$\frac{X_2 - X_3}{(X_2 + X_3)\ /2}$$

The nomograph devised for this triplicate test is also included in Figure 17.

In order to obtain triplicate data that could be legitimately compared with the quadruple data, one strength value was chosen at random from every quadruple and then discarded. The CV line for the "triplicate" sets was then determined as indicated by the solid line in Figure 17. The scales used in both nomographs in this figure were such that if both sets of data had the same CV, then their CV lines would fall on top of each other. As can be seen, they do not exactly coincide but fall very close to each other. The CV's determined for the triplicates and quadruples are .050 and .053 respectively. Hence, it would seem that the extra precision gained by using four specimens instead of three is not worthwhile. This would especially be true in large investigations involving many hundreds of test conditions.

Method of selecting specimens so as to minimize inherent specimen differences due to time or other factors

Another factor which enters into unconfined compressive strength testing involves the selection of specimens in order to equalize inherent specimen differences due to time or other factors. For example, in a typical soil-cement investigation, it may be necessary to determine the unconfined compressive strengths of a particular combination of soil and cement after 7, 28 and 120 days. In such a case, it is common practice to prepare nine specimens from the one batch of soil and cement and place them in the curing chamber together. Then at the end of 7 days, three samples are randomly chosen and tested, three others after 28 days and the remaining three after 120 days. It is a well-known fact that, as the specimens are being molded, the cement in the mixture is hydrating. Hence it is very possible-depending, of course, on the length of time it takes to prepare the specimens-that there may be significant differences between the last few specimens and those prepared at first; these differences may then be reflected in the strengths obtained after the specimens are tested. Oftentimes, random selection has the effect of equalizing these strength differences. On other occasions it does not do so. Certainly a method that is more reliable than chance is needed. In such cases, the following procedure is recommended.

1. Divide the nine specimens into three sets of three as indicated in Figure 18a. Call these sets, P, Q and R respectively. To each number within a set, assign a letter A, B or C as indicated.

2. Prepare a 3 x 3 "Latin Square" distribution for the letters A, B, C, as indicated in Figure 18b.

3. Taking note of the distribution in Figure 18b, select specimen A from set P, C from Q and B from R i. e. specimens number 1, 6 and 8. Then take B from set P, A from Q and C from R i. e. specimens 2, 4

Set	P	Q	R
Specimen Number	1 2 3	4 5 6	7 8 9
Letter	A B C	A B C	A B C

(a)

A	B	C
C	A	B
B	C	A

(b)

	Curing Time		
	7 day	28 day	120 day
Specimen Number	1	2	3
	6	4	5
	8	9	7
Sum of Specimen Numbers	15	15	15

(c)

	Curing Time		
	7 day	28 day	120 day
Specimen Number	1	3	2
	5	4	6
	9	8	7
Sum of Specimen Numbers	15	15	15

(d)

Figure 18. Combinations of specimen numbers that will minimize inherent specimen differences

and 9. Finally, the remaining specimens are C, B and A i. e. specimens 3, 5 and 7.

The final division of specimens is then as shown in Figure 18c. As can be seen, this method of selection is based upon the fact that if the sum of the specimen numbers is the same for a given curing period then the sum of their strengths should be the same. As a result, their averages should be the same.

By examining Figure 18b again, it is also obvious that another combination is possible. This combination is indicated in Figure 18d.

In an investigation where it may be necessary to have three curing periods and four specimens to be tested at each curing period, it is not possible to get a perfect distribution. Of the twelve specimens, it is not illogical to believe that the first three are the least subject to variability if a time trend exists. Hence, it would probably be best to distribute the last nine specimens in the above indicated manner and then randomly assign specimens 1, 2 or 3 to each of the obtained combinations.

The use of control specimens in detecting outliers

In the earlier part of this section a method was presented by which outliers could be detected in a series of soil-additive strength determinations involving four specimens per test condition. In Part 2 there was also presented a similar treatment for studies involving three specimens per test condition. Necessary to both of these methods was the establishment of a constant coefficient of variation for the investigation. One of the basic assumptions underlying these procedures is that the strength values used in calculating the CV come from a single normal population of values. This may not always be true-in fact it is very possible that, in certain studies, two or more normal populations may be involved. This, of course, means that there is more than one CV for the study and, hence, more than one critical disqualifying percentage. If this be so, then the obvious question arises as to how one can determine whether such divisions exist. The following practical example indicates one way of doing this.

Example To obtain data for this study, an investigation was undertaken involving many variables. As part of that study, some three hundred mixture-batches were prepared and three specimens were taken from each batch. Each batch was different from another by at least one of the following variables:

Soils - 2; a natural loess from western Iowa and an artificial mixture of sand and loess which, for reference sake, will be called the Colfax mix.

Cement type - 1: Type I Portland Cement

Cement contents - 3; 5, 8 and 11 percent.

Fly ash type - 3; each one from a different source.

Fly ash contents - 4; 0, 3, 6 and 9 percent

Moisture contents - 5; each moisture content was different for each combination of the other variables.

The specimens molded from these batches were all cured for seven days at the same relative humidity and temperature, before being tested in unconfined compression.

The preparation of these batches/specimens was routine, with the following exception. After every tenth batch was processed, a special batch-hereafter called a "control" batch-was prepared. Each of these control batches contained exactly the same amount of ingredients of the same cement, soil and water. Three specimens were prepared from each control batch, by the same operator, using the same compaction apparatus, procedures, etc. In all, twenty-seven of these control batches were prepared. The main reason for the preparation of these control batches, and hence the control specimens, was the feeling that if a constant CV did exist for the series as a whole, then certainly it would be reflected in the results obtained from the control specimens. Then, ideally, if the assumptions of one CV and one population are correct, the CV line for the main study should coincide with the CV line for the control specimens.

Using the afore mentioned procedures for triplicates the control specimen values and the main study values were plotted on the Outlier Paper as indicated in Figure 19. The values and calculations upon which main study graph is based are shown in Appendix C and those for

Table 4. Data for control specimens that were tested in unconfined compression after 7 days moist curing and 1 day immersion.

Sub-batch number	Individual strengths, psi	Average strength, psi ($\bar{X}$)	Range (R)	$r = \frac{R}{\bar{X}}$	% of other r values less than the given r value + 1/2 n
1	1145 1142 1063	1117	82	.0734	49.99
2	1191 1040 1022	1084	169	.1559	98.15
3	1178 1142 1135	1152	43	.0373	35.18
4	1168 1155 1135	1153	33	.0286	24.07
5	1254 1149 1145	1182	109	.0922	64.81
6	1093 1079 1060	1077	33	.0306	27.78
7	1109 1109 1093	1101	16	.0145	5.55
8	1303 1227 1148	1226	155	.1264	79.63
9	1326 1208 1201	1245	125	.1004	72.22
10	1231 1181 1162	1191	69	.0579	46.29

Table 4 (Continued).

Sub-batch number	Individual strengths, psi	Average strength, psi ($\bar{X}$)	Range (R)	$r = \frac{R}{\bar{X}}$	% of other values less than the given r value + 1/2 n
11	1221 1218 1181	1206	40	.0332	31.48
12	1185 1096 1073	1118	112	.1002	68.52
13	1106 1102 1099	1102	7	.0064	1.85
14	1273 1254 1099	1209	174	.1439	90.73
15	1135 1102 974	1071	161	.1503	94.44
16	1073 1052 1047	1057	26	.0246	16.66
17	1152 1135 1106	1131	46	.0407	38.88
18	1024 1020 1002	1015	22	.0217	12.96
19	1180 1079 1036	1099	144	.1310	83.34
20	1224 1086 1063	1124	161	.1432	87.03

Table 4 (Continued).

Sub-batch number	Individual strengths, psi	Average strength, psi ($\bar{X}$)	Range (R)	$r = \frac{R}{\bar{X}}$	% of other values less than the given r value + 1/2 n
21	1310 1300 1277	1295	33	.0255	20.37
22	1290 1218 1191	1233	99	.0803	53.70
23	1399 1389 1376	1388	23	.0166	9.26
24	1288 1270 1185	1248	103	.0825	57.40
25	1221 1122 1096	1146	125	.1091	75.92
26	1293 1244 1237	1258	56	.0445	42.59
27	1280 1198 1171	1216	109	.0896	61.11

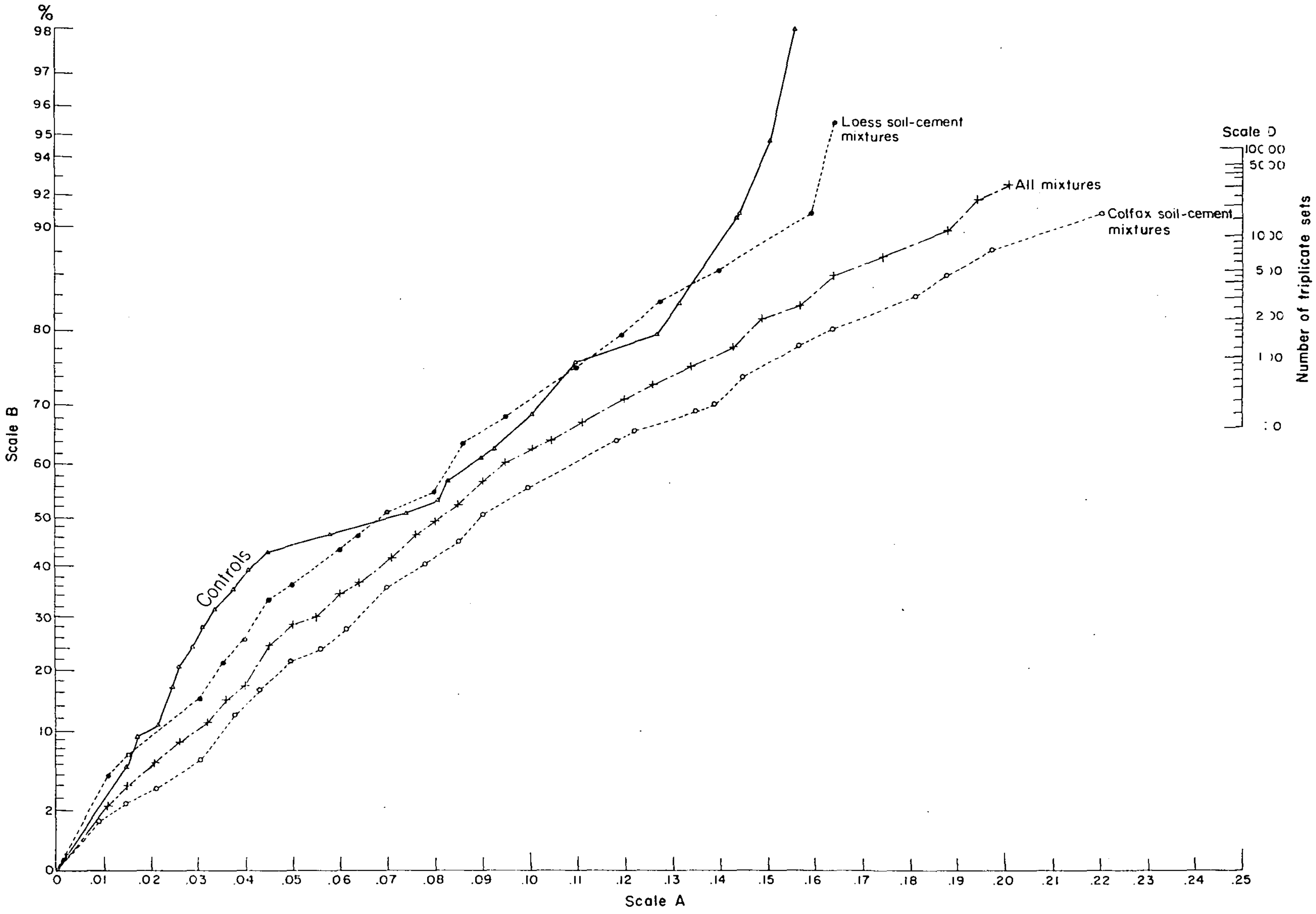
the control plot are shown in Table 4. The control specimen values form a rather irregular line due to the paucity of data. Hence it is rather difficult to estimate exactly where the CV line for the controls is located. However, one factor is quite clear-that there is little relationship between the control line and the modified cumulative distribution line (CV) obtained when all the mixture values were plotted. This automatically leads to the conclusions that the control CV line is not the same as the CV line that would be obtained from all the data.

In endeavoring to discover the exact cause of this non-coincidence, one of the first thoughts was that, perhaps, the difference between the soils was being reflected. As a result, the data was divided on the basis of soil type and replotted. These plots are shown in Figure 19 also. It is quite clear that this division has a significant effect on cumulative distribution lines. That plotted for the loess mixtures showed a definite upward movement to fall very close to the controls' cumulative distribution line, while the Colfax mixture results showed a definite movement in the opposite direction. However, this cannot be assumed to be the sought-after division, as the control specimens were all prepared from mixtures containing the Colfax soil-and not the loess soil.

In order to find the correct line of division, the data was further divided in many ways, such as type of fly ash, fly ash content, etc. While all of these caused minor changes, none was accepted as being adequate. Finally the data was divided in the following manner, and this is believed to be the correct division.

There is a phenomenon, well known in soil engineering, called the Moisture-Density Relationship (22). If a given amount of moisture is contained in a soil mixture and a given compactive energy is applied to that mixture, a certain density-usually expressed in pounds of dry soil per cubic foot-is obtained. If a little more moisture is added to the mixture and the same compactive energy is applied, the dry density will usually increase. As more and more moisture is added, and the same compactive effort applied each time, the dry density will keep increasing until a maximum value has been obtained, after which it will start to decrease. Now, using the moisture-density curves obtained

Figure 19. Illustration of the use of control specimens in determining if a constant CV exists for a series of strength determinations.


%
98
97
96
95
94
92
90
80
70
60
50
40
30
20
10
2
0
Scale B
0
.01
.02
.03
.04
.05
.06
.07
.08
.09
.10
.11
.12
.13
.14
.15
.16
.17
.18
.19
.20
.21
.22
.23
.24
.25
Scale A
Controls
Loess soil-cement mixtures
All mixtures
Colfax soil-cement mixtures
Scale D
10000
5000
1000
500
200
100
10
Number of triplicate sets

for every combination of soil, cement and fly ash, the strength values were divided on the following basis: those that were obtained from specimens compacted at moisture contents at or below optimum moisture content for maximum dry density were placed in one population and those compacted at moisture contents above the optimum were placed in another. By pure coincidence, it turned out that very close to half the specimens were in each category. These data were then plotted on the Outlier Paper and are shown in Figure 20.

It is quite evident in this figure that this separation is very valid. There is a definite division between the cumulative distribution lines for the below optimum strength values and the above optimum ones. The respective CV's as a result are .05 and .08. As a further check, the data was again subdivided on the basis of the other variables, but the resulting changes in the CV's were deemed insignificant. The further subdivision on the basis of soil type is shown in Figure 20 also. Using the estimated CV lines, a disqualifying percentage of 11.2 was found for the below optimum data and 17.6 for the above optimum data. These values caused 5 sets to be suspected from the below optimum data ones and 16 sets from the above optimum ones. After elimination of these suspect sets, the remaining ones were replotted on the outlier paper. These plots are shown in Figure 21. In this figure, both CV lines show definite straightening tendencies. Particularly is this noticeable with the below optimum values as the line entwines itself about the control distribution line. The above optimum values also show this straightening tendency but make little effort to align themselves with the controls.

The use of control specimens in evaluating the uniformity of materials during the investigation

One of the main causes of conflicting data is many large soil engineering investigations is believed to be that of inadequate preparation of the soil sample prior to testing. After the soil has been carried to the laboratory from the field, it is of course axiomatic that it should be thoroughly mixed before being used. Oftentimes, investigations may

Figure 20. Nomographic computation of the CV for specimens divided on the basis of their being above or below optimum moisture content for maximum density.

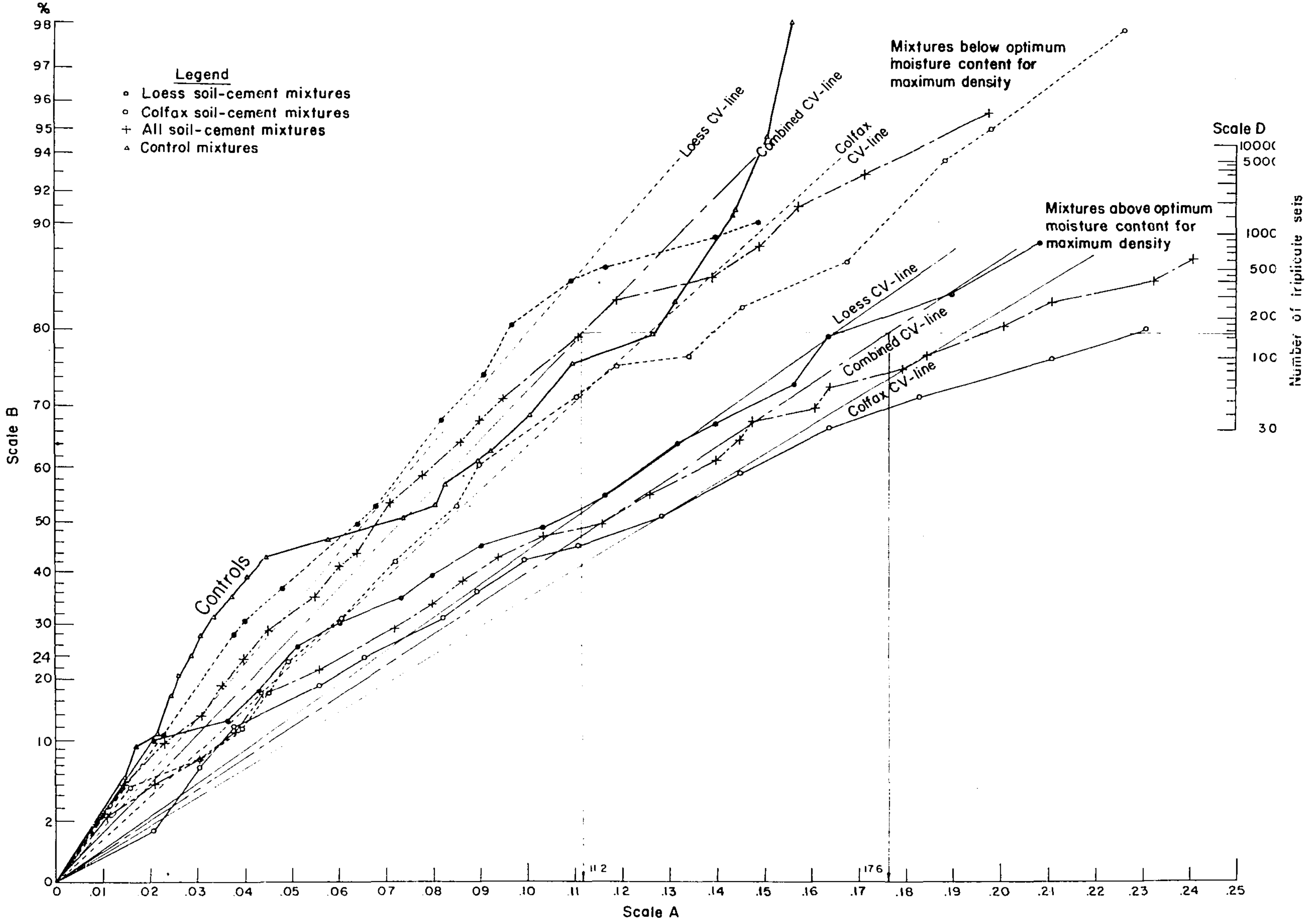
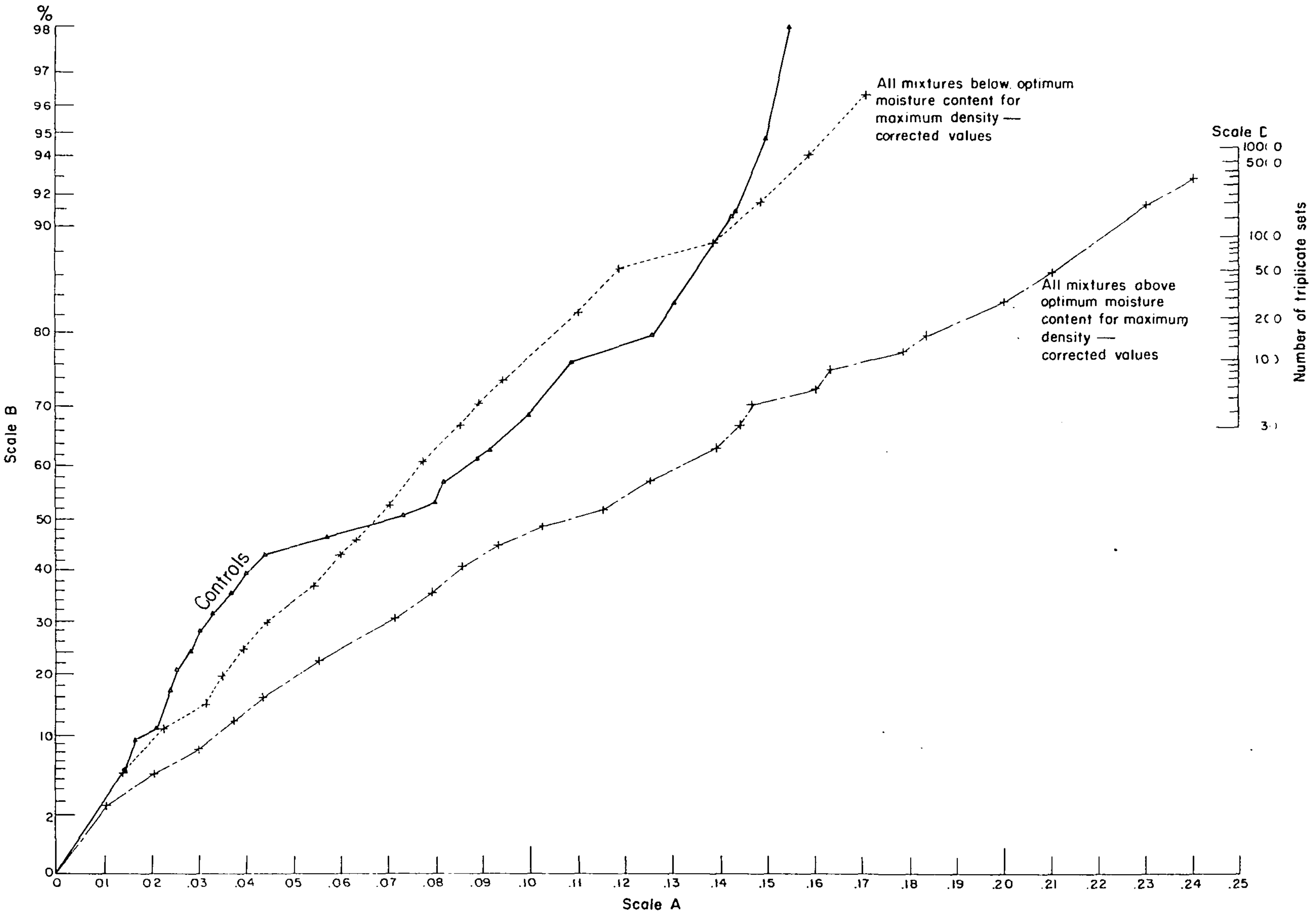

%
Legend
Loess soil-cement mixtures
Colfax soil-cement mixtures
All soil-cement mixtures
Control mixtures
Loess CV-line
Combined CV-line
Colfax CV-line
Controls
Mixtures below optimum moisture content for maximum density
Mixtures above optimum moisture content for maximum density
Scale D
Number of triplicate sets
10000
5000
1000
500
200
100
30
Scale B
98
97
96
95
94
92
90
80
70
60
50
40
30
24
20
10
2
0
Scale A
0
.01
.02
.03
.04
.05
.06
.07
.08
.09
.10
.11
.12
.13
.14
.15
.16
.17
.18
.19
.20
.21
.22
.23
.24
.25
112
176

Figure 21. Corrected nomographic computation of the CV for specimens divided on the basis of their being above or below optimum moisture content for maximum density.


%
98
97
96
95
94
92
90
80
70
60
50
40
30
20
10
2
0
Scale B
0
.01
.02
.03
.04
.05
.06
.07
.08
.09
.10
.11
.12
.13
.14
.15
.16
.17
.18
.19
.20
.21
.22
.23
.24
.25
Scale A
Controls
All mixtures below optimum moisture content for maximum density — corrected values
All mixtures above optimum moisture content for maximum density — corrected values
Scale C
1000.0
500.0
100.0
50.0
20.0
10.0
3.0
Number of triplicate sets

be carried on over the period of many months-perhaps even years. Over these long periods of time, differential settlement of particles may take place within the stored soil sample. Thus, the soil that is used at the beginning of the study may not be the "same" as that used towards the end. To illustrate this, the following example is cited.

Example The discussion previous to this illustrated the use of controls in detecting outliers. One of the soils used in that study was an artificial mixture of sand and loess known as the Colfax mix. This material had originally been thoroughly mixed but, at the time of this investigation, had been lying undisturbed in a bin for approximately 1 1/2 years. As it was suspected that differential settlement might have taken place within the mixture, the sand-loess mass was not given a thorough re-mixing before being used. As already mentioned, control batches containing the same ingredients were prepared during the course of the investigation. In all, twenty-seven of these batches were prepared, from each of which came three specimens. Figure 22a is a plot of the average densities obtained with these controls versus the batch numbers from which these densities came. Figure 22b is a plot of strengths versus control numbers for the same specimens. The greater the control batch number, the farther down the bin the material for that particular batch was obtained. In both figures, there are definite fluctuations that seem to occur in cycles rather than chance. Since the same apparatus was used in preparing the specimens and since the one operator was judged to be skilled at his work, the most logical conclusion is that, in this case, these cyclic differences are due to material non-uniformity.

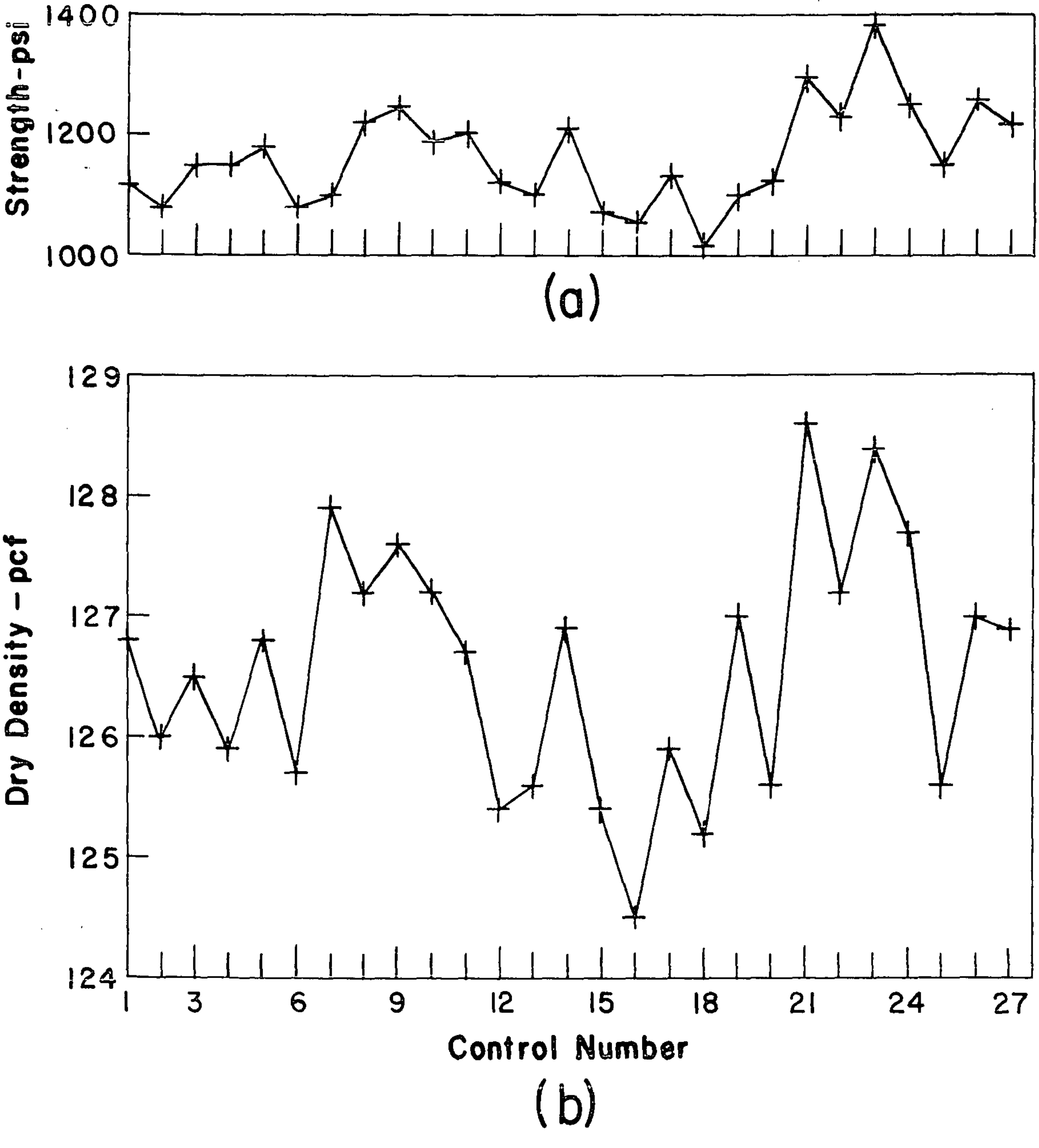
To avoid these material differences, the following recommendations are given.

1. Sample sufficient soil as is believed will be needed for the entire specific investigation.

2. After the original large soil sample has been brought to the laboratory and pulverized, sieved, etc. as required, it should be very thoroughly mixed.

3. After this thorough mixing, the soil sample should then be randomly divided up into batches. Each batch should not contain more

Figure 22. Illustration of the use of control specimens in detecting non-uniformity of experimental materials.


Strength-psi
1400
1200
1000
(a)
Dry Density - pcf
129
128
127
126
125
124
1
3
6
9
12
15
18
21
24
27
Control Number
(b)

than 30 to 40 pounds of material. These batches should then be boxed, covered and laid aside until needed.

4. As the study proceeds and soil is required, a box should be selected at random and the material in it should be thoroughly re-mixed before being used.

Since many stabilizing additives to soils may differ within themselves depending on their source etc. -it is axiomatic that a sufficient amount of the required additive should be at hand for the entire investigation. This additive should also be thoroughly mixed, batched and placed in sealed containers. If the additive is such that time affects its potency, this must be taken into account. Control specimens of the type already discussed can very easily be used to detect any such potency change.

In certain cases, having taken care of as many other areas of variability as possible, it may be desirable to check on the efficiency of the investigation operator. This can also be very easily done with the aid of the control specimens. A plot, then, of densities versus batch numbers, or strengths versus batch numbers, will clearly bring out any consistent erraticity of the operator.

Part 4 - Detecting Outliers in a Small Series of Soil-Additive Strength Determinations

Oftentimes, when in need of some specific information, the soil engineering investigator may feel it necessary to prepare only a few mixture-batches. Two problems face him in this situation. First of all, how many specimens should he prepare and test per batch, and secondly-as soon as he has obtained the strength values-what criteria should he use to detect outliers.

The problem of the number of samples is not a new one. Basically, the more specimens tested, the more reliable the results. From a practical viewpoint, the choice of number of samples is dependent upon the degree of accuracy required. This problem has been extensively treated elsewhere and so is not further discussed here. Instead, reference is made to ASTM Designation: E122-56, which presents the recommended practice for "Choice of Sample Size to Estimate its

Average Quality of a Lot or Process (23).

The problem of detecting outliers in this type of situation has also been extensively studied. Reference is made here to ASTM Designation: E178-61T which presents the recommended practice for "Dealing with Outlying Observations" (23). In this reference, two test criteria-each involving the use of the standard deviation-are presented. The first of these test criteria is as follows:

$$T_n = \frac{X_n - \bar{X}}{s}$$

where T_n = a test statistic

X_n = the highest and most suspect strength observation

$\bar{X}$ = arithmetic average of all n observations

s = estimate of the population standard deviation based on the sample data and calculated as follows:

$$s = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$$

If X_1, the lowest value, rather than X_n, is the doubtful value, the criterion is as follows:

$$T_1 = \frac{\bar{X} - X_1}{s}$$

The critical values for either case, for the 1 percent and 5 percent significance levels, are given in a table in that reference.

A second criterion is also given for detecting outliers

$$T'_1 = \frac{\bar{X} - X_1}{\sigma}$$

or

$$T'_n = \frac{X_n - \bar{X}}{\sigma}$$

This statistic is similar to the other except that σ is the known standard deviation as determined from independent sources, whereas s is an estimate of the standard deviation as determined from the present data. The critical values for T_1' and T_n' for the 5 percent and 1 percent significance levels, are also given in other tables in that reference.

Of these two criteria, obviously more confidence can be placed in the test criterion involving the known standard deviation, σ, as the estimate, s, is itself subject to contamination due to any possible outliers. However, from the point of view of the soil engineer, this is of little help as it is rarely in his work that the "true" standard deviation is known.

Purpose of the study

The purpose of this phase of the investigation was therefore to develop a more reliable criterion whereby invalid specimens in a small series of strength determinations involving the unconfined compressive strength test could readily and reliably be detected.

Proposed disqualification test

Three investigations have already been discussed intensively in this text, and, upon being examined, none of these three was found to have this constant standard deviation. Two of the studies-soil-calcium lignosulfonate (17) and soil-cement-fly ash-appear to have the essentially same coefficient of variation. The third, on soil-sodium silicate stabilization (18), had a different CV, but-as has been discussed earlier-this can perhaps be excused on the grounds of not being what might be loosely called a "normal" study. Another set of data (24) was also extensively examined and this turned out to have the same CV as the other two.

Based on this data, it appears that a constant CV exists for soil stabilization studies involving the unconfined compressive strength test and specimens compacted at or below optimum moisture content by the Iowa Compaction Apparatus. This CV appears to be equal to .050.

It is a known statistical fact that $\sigma = \bar{X}$. CV. Utilizing this fact, it is now possible to propose the following alternate, but powerful, outlier statistic for use in such soil stabilization studies:

$$T''_n = \frac{X_n - \bar{X}}{\bar{X}.\ CV}$$

or

$$T''_1 = \frac{\bar{X} - X_1}{\bar{X}.\ CV}$$

The critical values for T''_1 and T''_n for the 5 percent and 1 percent significance levels are given in Table 5.

An example of the use of this outlier test is now given.

Example involving the use of the proposed procedure As an illustration of the use of T''_1 and Table 5, consider the following four strength observations obtained during the course of the soil-cement-fly ash investigation: 678, 649, 625 and 540 psi. The doubtful value is $X_1 = 540$ psi. Then

$$\bar{X} = \frac{678 + 649 + 625 + 540}{4} = 637$$

$$CV = .050$$

Therefore

$$T''_1 = \frac{\bar{X} - X_1}{\bar{X}.\ CV} = \frac{637 - 540}{(637)(.050)} = \frac{97}{31.85} = 3.046$$

From Table 5, for n = 4 it can be seen that T''_1 as large as 2.16 would occur by chance with probability less than 0.05. In fact, for this particular illustration, it is clear that a T''_1 as large as 2.62 would occur by chance with probability somewhat less than 0.01. Thus the weight of the evidence is against the doubtful value as having come from the same normally distributed population as the other three.

Table 5. Critical values[a] of T''_1 and T''_n when the coefficient of variation (CV) is known

Number of observations	At 5 percent significance level	At 1 percent significance level
3	1.95	2.40
4	2.16	2.62
5	2.30	2.76
6	2.41	2.87
7	2.49	2.95
8	2.56	3.02
9	2.61	3.07
10	2.66	3.12
11	2.70	3.16
12	2.74	3.20
13	2.78	3.23
14	2.81	3.26
15	2.84	3.29
16	2.87	3.31
17	2.89	3.33
18	2.91	3.36
19	2.94	3.38
20	2.96	3.39
21	2.97	3.41
22	2.99	3.43
23	3.01	3.44
24	3.02	3.45
25	3.04	3.47

[a] The critical values presented in this table are for the known σ condition and are excerpted from ASTM Designation: E178-61T (23).

Discussion

Two factors are worth noting about the proposed outlier statistic T_1'' or T_n''. Firstly, it is a more reliable statistic than $T_n = (X_n - \bar{X})/s$ or its alternative T_1. The estimated standard deviation "s" is itself subject to contamination as it is determined on the basis of all the observations, including any possible outliers. On the other hand, the new statistic is not subject to such an error as it utilizes the known coefficient of variation, which is determined independently. Secondly, the proposed statistic has an overwhelming advantage over the other in terms of ease of computation, as there is no troublesome calculation of the standard deviation.

In certain instances involving a rather large number of strength observations, it is very possible that two or more outliers may occur on the same side as $\bar{X}$. In such an instance, a practical expedient is first to apply the test criterion to the innermost outlier while dropping, temporarily, the other outlier(s). If this test leads to the rejection of the innermost outlier, then the others are automatically rejected with it. If the innermost outlier is not rejected, the same procedure is then reapplied on the next potential outlier. In this manner the outliers are detected, as it were, from the inside out. It should be noted, however, that the theoretical basis of the test is somewhat violated by this procedure; in practice, the effect will generally not be significant. Nevertheless, it will probably be better to use a lower significance level, as for instance, 1 percent instead of 5 percent.

In a situation involving outliers on both sides of $\bar{X}$, again a practical expedient is to use the test first on one side and then on the other, in each case dropping temporarily the outlier on the opposite side. In this situation, it is again recommended that a lower significance level, say 1 percent, be used.

Finally, note should be taken of the fact that the coefficient of variation, as determined from the nomographic procedures described in Parts 2 and 3, is only accurate for values less than 0.15. This, therefore, indicates that the procedure recommended in this section is only adequate for CV values less than 0.15.

Part 5 - Method for Evaluating the Reliability of a Curing Chamber and Any Operator Variability Due to Time Trends

In many investigations, soil-additive specimens have to be cured for long periods of time before being tested. For comparison purposes, they are generally placed in a curing room where the temperature and relative humidity are kept at constant values. It is obvious that the temperature and humidity should be checked at regular intervals-at least once a day-and any marked variations noted and taken into account when the data is being evaluated. This, of course, is usually done in most laboratories. One check that is often forgotten, however, is that significant temperature and humidity differences may occur within the curing room itself, e.g. from top shelf to bottom shelf or from front to back, etc.

Purpose of the study

The purpose of this study was to develop and present a method whereby reflections of any differences due to temperature and/or humidity within the curing room could be detected. Since, on most occasions, specimens are placed in a curing room so as to gain strength under controlled conditions, it was decided to use a strength criterion to detect any possible differential effects. In addition, a method is given for determining if there is any significant time trend being reflected as specimens are being prepared for testing.

Proposed test procedure

Control specimens can be used to detect any such differences. However, control specimens as specified in this instance are slightly different than heretofore. For this case, it is recommended that-at a particular time-three separate batches be prepared and one specimen taken from each batch. In other words when reference is now made to specimen 1a, it means the specimen taken from batch "a" at time period one, and 1b means the specimen prepared from the second batch, "b", prepared at the same time period, etc.

For a proper evaluation, care has to be taken that the internal differences within the control specimens do not predominate and thus cloak any possible curing room positional effects. This necessarily involves very careful location of specimens along the curing room wall. The manner in which this is done is illustrated in Figure 23. This necessitates the total preparation of 27 batches, prepared at 9 different time periods. Thus, the number 4b on the diagram indicates the location on the curing room wall of the specimen prepared from the second batch at the fourth time period.

The most important feature of the proposed test is this location of the specimens in the curing room. Examination of this diagram indicates a perfectly balanced arrangement. The specimens are so arranged that the sum of the specimen numbers in any plane is equal to the sum of the specimen numbers in any other plane. Thus theoretically, it can be assumed that, if no curing room differences exist, the strengths on any plane should be equal to-within, of course, expected sampling differences-the strengths on any other plane.

On the basis of the above, it is now possible-by virtue of the complete cross-balancing (or orthogonality) achieved by the experimental design-to perform valid F-tests in which the "numerator sums of squares" are the usual simple indices of effect. These simple F-tests for the curing room are as follows:

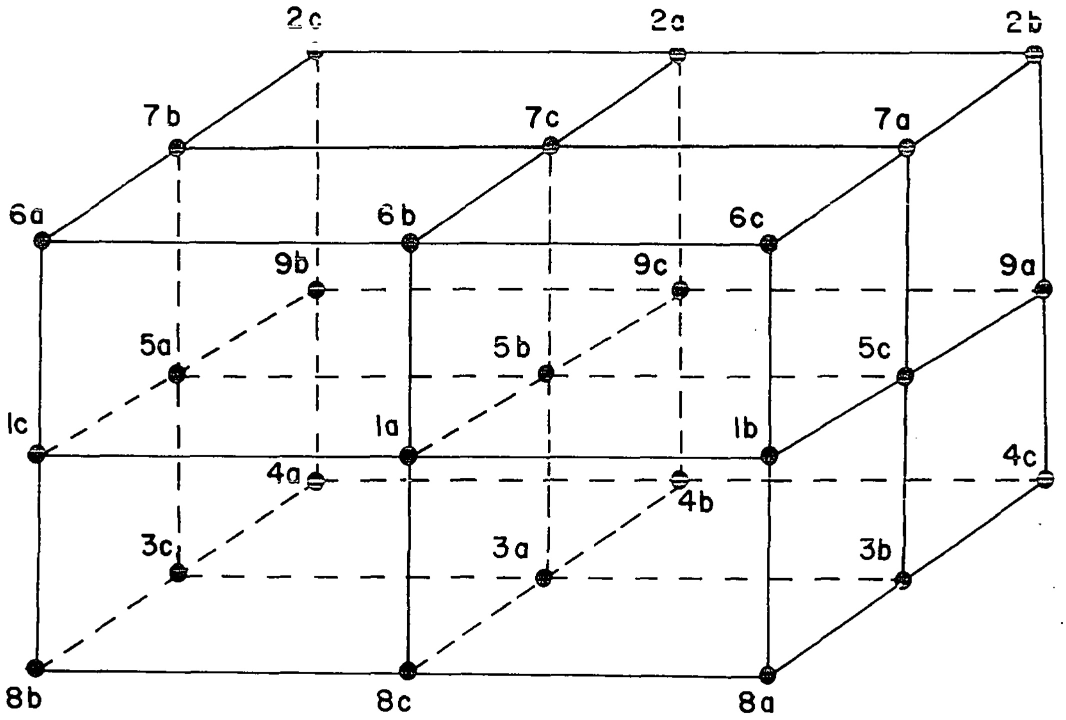
$$F_{\text{Height}} = \frac{\text{Mean square for height}}{\text{Error mean square}}$$

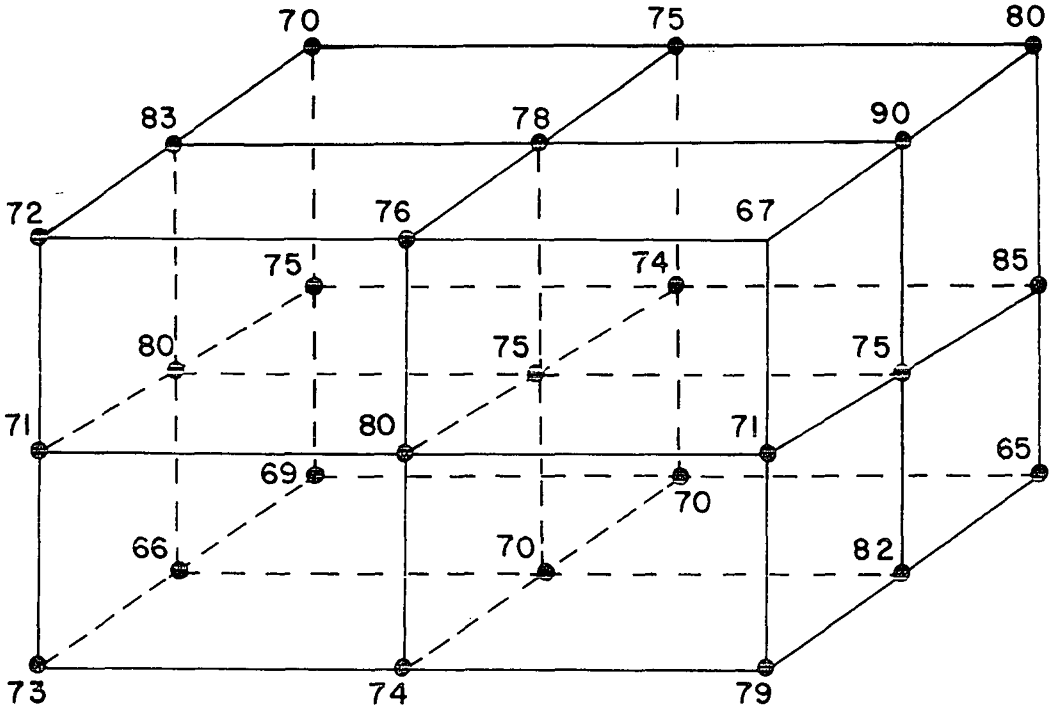
$$F_{\text{Width}} = \frac{\text{Mean square for width}}{\text{Error mean square}}$$

$$F_{\text{Depth}} = \frac{\text{Mean square for depth}}{\text{Error mean square}}$$

Figure 23. Location of specimens along the curing room wall.

Figure 24. Location of strength values along the curing room wall.


2c
2a
2b
7b
7c
7a
6a
6b
6c
9b
9c
9a
5a
5b
5c
lc
la
lb
4a
4b
4c
3c
3a
3b
8b
8c
8a


70
75
80
83
78
90
72
76
67
75
74
85
80
75
75
71
80
71
65
69
70
66
70
82
73
74
79

While performing these checks, it may also be desirable to check on any consistent errors due to time trends. These can be of two types. The first involves testing for any consistent differents between specimens made from batches prepared at the same time, e.g. between specimens a, b and c. The test for this is:

$$F_{Short} = \frac{\text{Mean square for short time}}{\text{Error mean square}}$$

Long term strength differences, as reflected by possible differences between specimens made from batches prepared a long time apart, i.e. consistent differences between batches prepared at time period one versus time period two versus time period three, etc., can be detected by the following test:

$$F_{Long} = \frac{\text{Single degree-of-freedom mean square for time}}{\text{Error mean square}}$$

An example involving the use of the proposed procedure is now given.

Example involving the use of the proposed procedure

To illustrate the method, mythical strength values were assumed. The locations to which these mythical values were assigned are shown in Figure 24.

Compute

$$MS_{Height} = \frac{\dfrac{\sum_{i=1}^{3} X_i^2}{9} - \dfrac{X^2}{27}}{2}$$

where X_i = sum of all 9 strength values obtained at height "i", and

X = sum of all 27 strength values.

Thus

(sum of all values in the top plane)^2 = 691^2 = 477481

(sum of all values in the middle plane)^2 = 686^2 = 470596

(sum of all values in the bottom plane) = 648 = 419904

(sum of all 27 values)^2 = 2025^2 = 4100625

Therefore

$$MS_{Height} = \frac{\frac{1367981}{9} - \frac{4100625}{27}}{2}$$

$$= 61.5$$

Similarly

$$MS_{Width} = \frac{\frac{\sum_{i=1}^{3} X_i^2}{9} - \frac{X^2}{27}}{2}$$

$$= \frac{\frac{1367501}{9} - \frac{4100625}{27}}{2} = \frac{\frac{1367501}{9} - \frac{4100625}{27}}{2}$$

$$= 34.75$$

and

$$MS_{Depth} = \frac{\sum_{i=1}^{3} X_i^2 - \frac{X^2}{27}}{9}$$

$$= \frac{\frac{1367739}{9} - \frac{4100625}{27}}{2}$$

$$= 48.00$$

Compute

$$MS_{Short} = \frac{\frac{\sum_{i=1}^{3} X_i^2}{9} - \frac{X^2}{27}}{2}$$

where X_i = sum of all 9 strength values obtained from "i" batches,

and X = sum of all 27 strength values.

Thus

$$(\text{sum of all values from "a" batches})^2 = 700^2 = 490000$$

$$(\text{sum of all values from "b" batches})^2 = 685^2 = 469225$$

$$(\text{sum of all values from "c" batches})^2 = 640^2 = 409600$$

Therefore

$$MS_{Short} = \frac{\frac{1,368825}{9} - \frac{4100625}{27}}{2}$$

$$= 70.80$$

Compute

$$MS_{Long} = \frac{\frac{\left[\sum\limits_{i=1}^{9} (T_i)(X_i)\right]^2}{(3)(\Sigma T_i^2)}}{1}$$

where

X_i = sum of all three values obtained at time period "i", and

T_i = a coded time factor, varying from -4 to +4,

Thus

$$\sum_{i=1}^{9} (T_i)(X_i) = \Big[(-4)(80+71+70)+(-3)(75+80+70)+(-2)(70+82+66) + (-1)(69+70+65)+(0)(80+75+75)+(1)(72+76+67) +(2)(90+83+78)+(3)(79+73+74)+(4)(85+75+74)\Big]^2$$

$$= [128]^2$$

$$= 16384$$

and

$$(3)(\Sigma T_i^2) = (3)\left[(-4)^2+(-3)^2+(-2)^2+(-1)^2+(0)^2+(1)^2+(2)^2+(3)^2+(4)^2\right]$$

$$= 180$$

Thus

$$MS_{Long} = \frac{16384}{180} = 91.02$$

Compute

$$MS_{Error} = \frac{\Sigma(Y-\bar{Y})^2 - SS_{Height} - SS_{Width} - SS_{Depth} - SS_{Short} - SS_{Long}}{17}$$

where Y = individual strength value,

$\bar{Y}$ = mean of all 27 strength values,

SS_{Height} = sum of squares due to height

= (MS_{Height})(degrees of freedom associated with it)

and similarly

$SS_{Width} = (MS_{Width})(df)$

$SS_{Depth} = (MS_{Depth})(df)$

$SS_{Short} = (MS_{Short})(df)$

$SS_{Long} = (MS_{Long})(df)$

Thus

$$\Sigma(Y-\bar{Y})^2 = (80-75)^2 + (71-75)^2 + \ldots\ldots + (74-75)^2$$

$$= 642$$

Therefore

$$MS_{Error} = \frac{642-(61.50)(2)-(34.75)(2)-(48.00)(2)-(70.80)(2)-(91.02)(1)}{17}$$

$$= 7.11$$

F-test,

$$F_{Height} = \frac{61.50}{7.11} = 8.66$$

$$F_{Width} = \frac{34.75}{7.11} = 4.89$$

$$F_{Depth} = \frac{48.00}{7.11} = 6.76$$

$$F_{Short} = \frac{70.80}{7.11} = 9.97$$

From tabulated values (25), $F_{2,17} = 3.59$ at the 95% level. All the above values are well above this value thus indicating that, for the assumed strength values, differences within the curing room and from batches a to b to c are significant.

F-test,

$$F_{Long} = \frac{91.022}{7.11} = 12.81$$

From tabulated values (25), $F_{1,17} = 4.45$ at the 95% level. Thus, it appears, that for the assumed strength values, strength differences due to long term time trends are significant.

Discussion

A few words should be said about some of the conditions upon which this test is based.

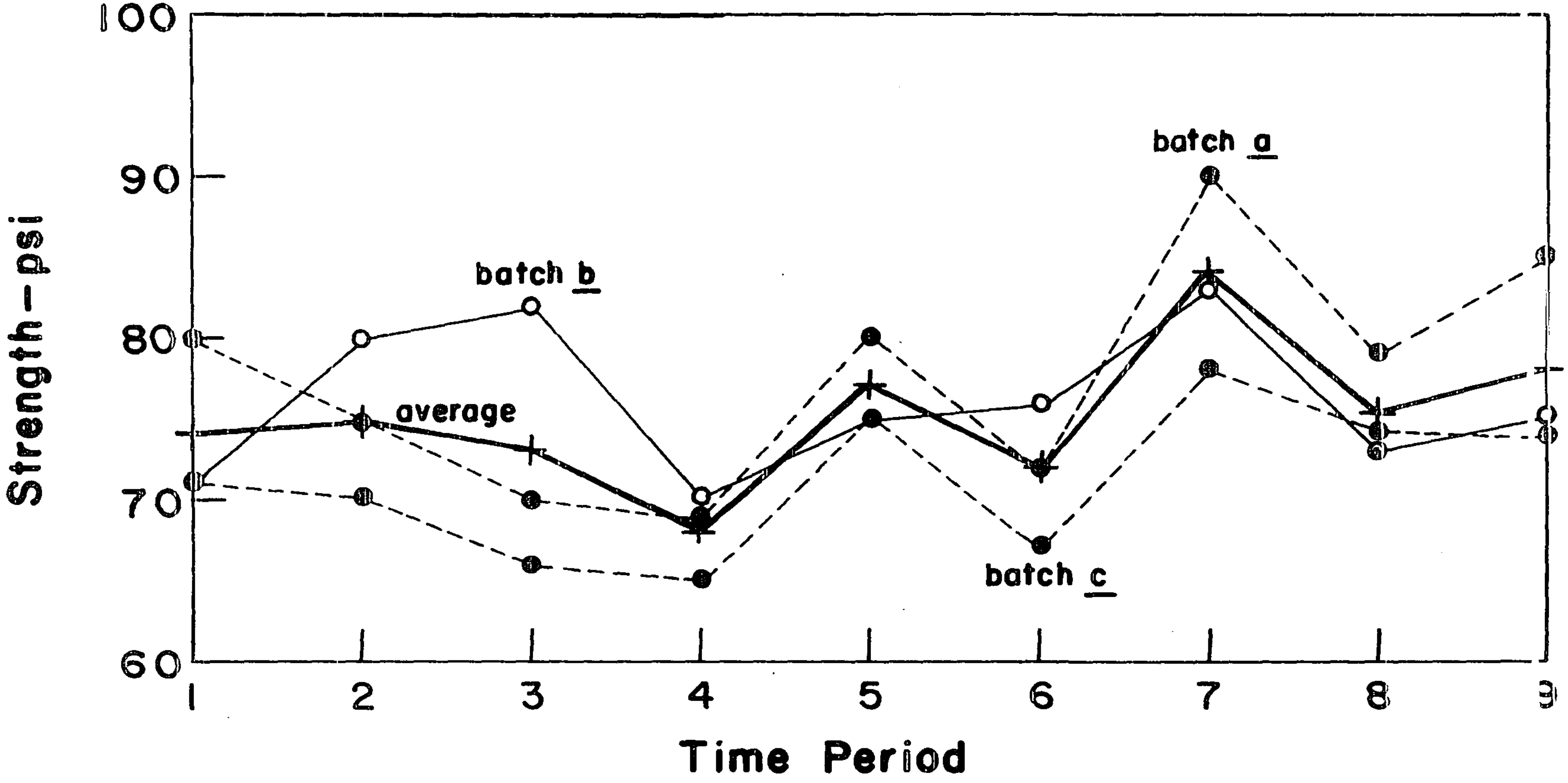
Firstly, the specimens must be arranged in balanced factorial form as indicated by Figure 23.

Secondly, this analysis is only applicable to situations involving homogenous error. In other words, only one type of error must be involved. For instance, in the above example, the only error allowed for was the batch to batch error for the 27 batches. Of course, this test would also be applicable to a situation involving 27 specimens where all 27 came from the one batch. In this situation, the only error allowed for would be the within-batch error, as there is now no between-batch error. Because of this requirement, this test is not applicable where the control specimens are prepared from the same batch as well as from different batches-as illustrated by the control specimens discussed in Part 3-as these strength values contain both a within-batch component of error and a between-batch component of error.

Thirdly, this test is only applicable to situations where any possible time trends are linear. If a time trend higher than linear-as illustrated by Figure 22 in Part 3-enters into the problem, non-orthogonality is created and this requires a more difficult analysis. The easiest way to check for non-linearity is to plot the mean of the strength values for each time period as a function of time and note-by eye-if there is non-linearity. The plot of the mythical values assumed in the example is shown in Figure 25. Since it appeared as if there were no curve-linear changes in this graph, linearity was assumed in the illustrated example.

It is worth noting at this stage that Figure 25 indicates very clearly what has been proved by the statistical analysis. There appears to be a long term time trend and this was shown to be so by the analysis.

Figure 25. Plot of mythical strength values as a function of time.


100
90
80
70
60
Strength – psi
batch b
average
batch a
batch c
1
2
3
4
5
6
7
8
9
Time Period

In addition-and more obviously-there is a short term time trend. This is very strongly indicated by the fact that the plot of specimens from the "a" batches is consistently higher than the plot of the "c" specimens, as both plots tend to be nearly parallel to each other.

SUMMARY AND CONCLUSIONS

The primary purpose of this investigation was to provide some statistical procedures which would help the soil engineering researcher to control and evaluate his results. The following procedures were therefore developed and presented:

1. Graphical method for detecting outliers in a typical correlation study. This method is applicable to studies where quadruples of specimens are prepared per test condition and where each member of the quadruple comes from a different batch.
2. Regression analysis for determining if a relation exists between two methods of testing a soil, when both methods of test are subject to error. This procedure is a "least squares" method that will give the same answer whether Y is thought of as being regressed on X or X on Y.
3. Graphical method for detecting outliers in a large series of soil-additive strength determinations involving triplicates of specimens. For this procedure, each triplicate set must come from the same batch.
4. Graphical method for determining the reliability-as a whole-of an investigation that involves large numbers of strength determinations. This method is only applicable to studies involving triplicates of specimens, where each triplicate set comes from the same batch.
5. Graphical method for detecting outliers in a large series of soil-additive strength determinations. This method is applicable to studies involving quadruples of specimens, where each quadruple set comes from the same batch.
6. Graphical method for determining the increase in precision when four specimens-instead of three-are used per test condition. In the example given, it was found that the extra precision gained by using four specimens was not worthwhile.

7. Method for selecting specimens for testing so as to minimize inherent specimen differences due to time or other factors. This method is most applicable when the numbers of specimens prepared at the one time are such that their square roots can be obtained, e.g. 9, 16 or 25 specimens.
8. Methods-involving the use of control specimens-for evaluating the validity of an investigation involving large numbers of soil-additive strength determinations. These control specimens can be used to determine if all the strength values come from the one population, and, also to detect any material or operator variability throughout the investigation.
9. Some recommendations regarding the preparation of a soil sample prior to the actual investigations. These recommendations are primarily aimed at eliminating material differences throughout the study itself.
10. Method for detecting outliers in a series of soil-additive strength determination involving small numbers of specimens. This method is only applicable to studies for which there is a known coefficient of variation.
11. Method for evaluating the reliability of a typical soil-additive curing chamber. This method is most useful in determining if there are reflections of temperature or humidity differences in various parts of the curing room itself. Until it is known whether such differences exist, it is recommended that all specimens for a particular investigation be cured in the same general area of the curing room.
12. Method for determining if there is any significant operator variability-due to a possible time trend-in a large series of soil additive strength determinations. This method is only applicable in situations where only homogeneous error is involved and where any possible time trends are linear.

The above procedures have been presented in as straightforward a manner as possible so that they may be useful to the soil engineering researcher who is not too familiar with statistical terminology. Although these methods have been presented under specific sub-titles and within specific situations, it is emphasized that they are intended to serve as prototypes for other similar types of soil engineering investigations.

FUTURE INVESTIGATION

The rapid growth of soil engineering-particularly soil stabilization-has opened new areas in which statistical procedures are needed and should be applied. To attempt to list all of these areas would be impossible. However, future investigation could be broadly divided into the following three phases:

1. Evaluation of the reliability and reproducability of the standard soil engineering tests.
2. Development of procedures useful in the design of soil engineering experiments.
3. Development of methods of evaluating the results obtained in soil engineering experiments.

It is emphasized that any procedures developed in Part 3 cannot really be utilized to their utmost until Parts 1 and 2 are thoroughly taken into account.

BIBLIOGRAPHY

Literature Cited

1. Shewhart, W. A. Some applications of statistical methods to the analysis of physical and engineering data. Bell System Technical Journal. 3: 43-87. 1924.

2. American Society for Testing Materials. 1958 book of ASTM standards. Part 4. Philadelphia, Pa. Author. 1958.

3. Davidson, D. T. and Handy, R. L. Lime and lime-pozzolan stabilization. In Woods, K. B., ed. Highway engineering handbook. pp. 21-98 to 21-123. New York, N. Y. McGraw-Hill Book Co., Inc. 1960.

4. American Society of Civil Engineers. Development of CBR flexible pavement design methods for airfields. Transactions, American Society of Civil Engineers. 115: 453-589. 1950.

5. Yoder, E. J. Principles of pavement design. New York, N. Y. John Wiley and Sons, Inc. 1959.

6. Maclean, D. S. Considerations affecting the design and construction of stabilized-soil road bases. J. of the Institution of Highway Engineers (London). 3, No. 9: 16-33. 1956.

7. American Society for Testing Materials. Committee D-18. Tentative method of test for determining the bearing ratio (CBR) of soils. (Mimeo.) Philadelphia, Pa. Author. 1959.

8. Davidson, D. T. and Chu, T. Y. Some laboratory tests for the evaluation of stabilized soils. Unpublished paper presented at the 58th annual meeting of the American Society for Testing Materials, Atlantic City, N. J., 1955. (Mimeo.) Ames, Iowa. Iowa State University of Science and Technology, Engineering Experiment Station. ca. 1955.

9. Hastings, C., Jr., Mosteller, F., Tukey, J. W. and Winsor, C. P. Low moments for small samples: a comparative study of order statistics. Annals of Mathematical Statistics 18: 413-26. 1947.

10. Miller, L. H. Table of percentage points of Kolmogorov statistics. J. of the American Statistical Association 51: 111-121. 1956.

11. Lindley, D. V. Regression lines and the linear functional relationship. J. of the Royal Statistical Society 9, Supplement: 235-237. 1947.

12. Williams, E. J. Regression analysis. New York, N. Y. John Wiley and Sons, Inc. 1959.

13. Aeton, F. S. Analysis of straight line data. New York, N. Y. John Wiley and Sons, Inc. 1959.

14. Waddell, J. J. Quality control in the construction industry. Industrial Quality Control 17, No. 7: 12-15. 1961.

15. Deming, W. E. Statistical adjustment of data. New York, N. Y. John Wiley and Sons, Inc. 1943.

16. Daniel, C. Use of half-normal plots in interpreting factorial two-level experiments. Technometrics 1: 333-335. 1959.

17. Demirel, T. Use of calcium lignosulfonate with aluminum sulphate for soil stabilization. Unpublished M. S. thesis. Ames, Iowa. Library, Iowa State University of Science and Technology. 1959.

18. Ruff, C. G. Sodium silicate-lime stabilization. Unpublished M. S. thesis. Ames, Iowa. Library, Iowa State University of Science and Technology. 1960.

19. Pearson, E. S. and Hartley, H. O. Biometrika tables for statisticians. Vol. 1. Cambridge, Mass. Cambridge University Press. 1956.

20. Nair, K. R. The distribution of the extreme deviate from the sample mean and its studentized form. Biometrika 35: 118-144. 1948.

21. O'Flaherty, C. S., Mateos, M. and Davidson, D. T. Effect of fly ash and sodium carbonate on soil-cement mixtures. Unpublished paper presented at the 41st annual meeting of the Highway Research Board, Washington, D. C., 1962. (Mimeo.) Ames, Iowa. Iowa State University of Science and Technology, Engineering Experiment Station. ca. 1962.

22. Proctor, R. R. Fundamental principles of soil compaction. Engineering News-Record 111, No. 9: 245-248. 1933.

23. American Society for Testing Materials. 1961 book of ASTM standards. Parts 3-11. Philadelphia, Pa. Author. 1961.

24. Remus, M. D. Strength of soil-lime mixtures at standard and modified A. A. S. H. O. densities. Unpublished M. S. thesis. Ames, Iowa. Library, Iowa State University of Science and Technology. 1959.

25. Ostle, Bernard. Statistics in research. Ames, Iowa. The Iowa State College Press. c1954.

Additional References

Anscombe, F. J. Rejection of outliers. Technometrics 2: 123-147. 1960.

Basu, D. On statistics independent of a complete sufficient statistic. Sankhya 15: 377-380. 1955.

Daniel, C. Locating outliers in factorial experiments. Technometrics 2: 149-156. 1960.

Dixon, W. J. Analysis of extreme values. Annals of Mathematical Statistics 21: 488-506. 1950.

Kruskal, W. H. Some remarks on wild observations. Technometrics 2: 1-3. 1960.

Lieblein, J. Properties of certain statistics involving the closest pair in a sample of three observations. J. of Research of the National Bureau of Standards 48: 255-268. 1952.

Mateos, M. Physical and mineralogical factors in stabilization of Iowa soils with lime and fly ash. Unpublished Ph. D. thesis. Ames, Iowa. Library, Iowa State University of Science and Technology. 1961.

Ruben, H. On the moments of order statistics in samples from normal populations. Biometrika 41: 200-227. 1954.

Youden, W. S. Sets of three measurements. Scientific Monthly 77: 143-147. 1953.

ACKNOWLEDGEMENTS

The author wishes to express his very sincere appreciation to Dr. D. T. Davidson, Professor of Civil Engineering, for his valuable guidance and assistance throughout this investigation-and, in fact, throughout the writer's entire time at the Soil Research Laboratory.

Special thanks are due to Dr. H. T. David of the Department of Statistics for his very patient, and generous, assistance and encouragement throughout this study.

The helpful advice given by Dr. R. L. Handy, Dr. M. Mateos, Mr. T. Demirel and Mr. J. Wang are gratefully acknowledged.

Last, but far from being least, very many sincere thanks are due to the author's wife, Nuala, who gave so much assistance "with heart an' hand".

APPENDIX A

Tabulation of data used in the correlation study

Cement content, %	Unconfined compressive strength values Sub-batch number	Strength, psi	Gap	Adjusted gap	Largest gap / 2nd largest gap (R_o)	% of other R_o's < given R_o	California Bearing Ratio values Sub-batch number	CBR, %	Gap	Adjusted gap	Largest gap / 2nd largest gap (R_o)
11	3 4 1 2	925 912 882 802	13 30 80	13 36.99 80	2.163	58	4 3 2 1	1140 1130 1097 990	10 33 107	10 40.686 107	2.630
	4 3 2 1	846 770 732 601	76 38 131	76 46.85 131	1.723	44	4 1 2 3	1295 1250 1085 1085	45 165 0	45 203.429 0	4.521
	1 2 3 4	797 747 350 329	50 397 21	61.65 397 21	6.440	93	2 1 3 4	1208 965 595 565	243 370 30	243 456.173 30	1.877
	1 2 3 4	769 756 579 547	13 177 32	13 218.22 32	6.819	95	2 4 1 3	1165 950 923 895	215 27 28	215 33.288 28	6.459
	2 4 1 3	610 563 425 410	47 138 15	47 170.14 15	3.620	80	2 4 1 3	726 700 677 638	26 23 39	26 28.357 39	1.375
	1 3 2 4	186 175 149 133	11 26 16	11 32.06 16	2.003	53	3 1 2 4	430 420 328 302	10 92 26	10 113.427 26	4.362
11	1 4 3	1181 1032 955	149 77 208	149 94.93 208	1.396	28	2 3 4	1280 1225 1180	55 45 120	55 55.31 120	2.163

	3	410	15	15			3	638	39	39	
	1	186	11	11			3	430	10	10	
	3	175	26	32.06	2.003	53	1	420	92	113.427	4.362
	2	149	16	16			2	328	26	26	
	4	133					4	302			
11	1	1181	149	149			2	1280	55	55	
	4	1032	77	94.93	1.396	28	3	1225	45	55.31	2.163
	3	955	208	208			4	1180	120	120	
	2	747					1	1060			
	2	994	94	94			3	1218	54	54	
	3	900	122	150.41	1.600	39	1	1164	34	41.919	1.288
	1	778	29	29			2	1130	40	40	
	4	749					4	1090			
	4	937	14	14			3	1300	74	74	
	2	923	33	40.69	3.662	81	1	1226	186	229.319	1.697
	3	890	149	149			2	1040	135	135	
	1	741					4	905			
	1	638	111	111			3	995	155	155	
	3	527	33	40.69	2.523	66	1	840	75	92.468	1.161
	2	494	44	44			4	765	180	180	
	4	450					2	585			
	1	489	202	202			1	700	95	95	
	3	287	2	2.47	9.619	97	4	605	0	0	5.588
	4	285	21	21			3	605	17	17	
	2	264					2	588			
	1	246	39	39			1	600	85	85	
	2	207	7	8.63	3.250	76	4	515	47	57.946	1.467
	3	200	12	2			3	468	53	53	
	4	188					2	415			
8	1	431	76	76			3	837	7	7	
	2	355	35	43.15	1.761	45	1	830	22	27.124	2.507
	3	320	7	7			4	808	68	68	
	4	313					2	740			
	1	394	40	40			2	760	40	40	
	3	354	32	39.45	1.014	2	4	720	33	40.386	1.017
	4	322	2	2			3	687	7	7	
	2	320					1	680			

8	1	431	76	76			3	837	7	7	
	2	355	35	43.15	1.761	45	1	830	22	27.124	2.507
	3	320	7	7			4	808	68	68	
	4	313					2	740			
	1	394	40	40			2	760	40	40	
	3	354	32	39.45	1.014	2	4	720	33	40.586	1.017
	4	322	2	2			3	687	7	7	
	2	320					1	680			
	3	392	3	3			4	612	16	16	
	4	389	8	9.86	2.230	60	2	596	11	13.562	5.937
	1	381	22	22			3	585	95	95	
	2	359					1	490			
	4	404	92	92			4	613	0	0	
	3	312	15	18.49	4.975	89	2	613	53	65.344	1.005
	2	297	16	16			3	560	65	65	
	1	281					1	495			
	3	300	37	37			3	567	15	15	
	4	263	37	45.62	1.233	18	4	552	52	64.111	1.336
	1	226	28	28			1	500	48	48	
	2	198					2	452			
	2	160	3	3			4	442	38	38	
	4	157	1	1.23	1.103	9	2	404	29	35.754	1.063
	3	156	6	6			1	375	18	18	
	1	150					3	357			
	4	630	70	70			3	980	40	40	
	2	560	48	59.180	1.183	14	1	940	10	12.329	2.250
	1	512	1	1			2	930	90	90	
	3	511					4	840			

---

dicates that the specimens were immersed for 24 hours before being tested.

Continued)

Cement ontent, %	Unconfined compressive strength values						California Bearing Ratio values				
	Sub-batch number	Strength, psi	Gap	Adjusted gap	Largest gap / 2nd largest gap (R_o)	% of other R_o's < given R_o	Sub-batch number	CBR, %	Gap	Adjusted gap	Largest gap / 2nd largest gap (R_o)
8	1	536	11	11			4	891	24	24	
	2	525	33	40.686	1.937	52	3	867	12	14.795	1.875
	4	492	21	21			2	855	45	45	
	3	471					1	810			
	2	545	25	25			1	940	70	70	
	3	520	20	24.658	3.600	80	3	870	90	110.961	1.585
	1	500	90	90			2	780	0	0	
	4	410					4	780			
	4	394	12	12			1	767	72	72	
	3	382	34	41.919	2.620	67	3	695	7	8.630	8.343
	1	348	16	16			2	688	8	8	
	2	332					4	680			
	2	382	31	31			3	667	15	15	
	1	351	20	24.658	1.257	20	2	652	2	2.166	2.667
	3	331	0	0			1	650	40	40	
	4	331					4	610			
	3	207	7	7			4	450	27	27	
	2	200	3	3.696	1.894	51	2	424	12	14.794	1.825
	4	197	3	3			1	412	11	11	
	1	194					3	401			
	3	137	7	7			2	295	14	14	
	4	130	2	2.466	2.143	57	1	281	21	25.891	1.849
	1	128	15	15			3	260	10	10	
	2	113					4	250			
5	3	171	19	19			4	367	32	32	
	4	152	17	20.959	1.103	9	1	335	27	33.288	1.292
	1	135	3	3			3	308	43	43	
	2	132					2	265			

	3 4 1 2	137 130 128 113	7 2 15	7 2.466 15	2.143	57	2 1 3 4	295 281 260 250	14 21 10	14 25.891 10	1.849
5	3 4 1 2	171 152 135 132	19 17 3	19 20.959 3	1.103	9	4 1 3 2	367 335 308 265	32 27 43	32 33.288 43	1.292
	1 2 3 4	164 164 160 130	0 4 30	0 4.932 30	6.083	92	2 3 1 4	370 358 330 300	12 28 30	12 34.521 30	1.151
	4 1 3 2	157 133 110 94	24 23 16	24 28.336 16	1.181	14	4 1 3 2	346 309 290 215	37 19 75	37 23.125 75	2.027
	3 4 1 2	144 141 90 63	3 51 27	3 62.878 27	2.329	62	2 3 4 1	320 310 302 217	Not taken into account for t outlier analysis		
	1 4 2 3	67 63 62 62	4 1 0	4 1.233 0	3.244	76	1 2 3 4	214 166 155 151	48 11 4	48 13.562 4	3.539
	3 4 2 1	184 175 167 156	9 8 11	9 9.863 11	1.115	9	1 2 3 4	452 419 419 400	33 0 19	33 0 19	1.737
	3 1 4 2	170 140 122 119	30 18 3	30 22.192 3	1.352	25	4 1 3 2	402 393 386 370	9 7 16	9 8.630 16	1.778
	4 2 1	148 139 134	9 5 42	9 6.165 42	4.667	88	1 2 4	345 345 322	0 23 92	0 28.357 92	3.244

3	184	9	9			1	452	33	33	
4	175	8	9.863	1.115	9	2	419	0	0	1.737
2	167	11	11			3	419	19	19	
1	156					4	400			
3	170	30	30			4	402	9	9	
1	140	18	22.192	1.352	25	1	393	7	8.630	1.778
4	122	3	3			3	386	16	16	
2	119					2	370			
4	148	9	9			1	345	0	0	
2	139	5	6.165	4.667	88	2	345	23	28.357	3.244
1	134	42	42			4	322	92	92	
3	92					3	230			
4	109	7	7			4	327	7	7	
3	102	2	2.466	2.428	64	3	320	20	24.658	1.121
1	100	17	17			1	300	22	22	
2	83					2	278			
3	80	5	5			3	249	3	3	
1	75	3	3.699	1.352	25	1	246	9	11.096	2.121
2	72	3	3			4	237	5	5	
4	69					2	232			
3	25	1	1			1	74	6	6	
4	24	1	1.233	1.233	18	3	68	5	6.165	1.027
2	23	0	0			2	63	0	0	
1	23					4	63			
1	41	2	2			1	128	0	0	
2	39	2	2.466	3.244	76	2	128	8	9.363	3.288
3	37	8	8			3	120	3	3	
4	29					4	117			

APPENDIX B

Table 7a. Tabulation of data used in determining the coefficient of variation (CV) for the study involving quadruples of strength determinations - Colfax soil

Cement content, %	Fly ash content, %	Fly ash content, Lab. no.	Sodium carbonate content, %	1 day immersed, 7 day moist cured, unconfined compressive strength values, Individual strengths, psi				Average strength, psi	$\bar{X}=(X_2+X_3)/2$	Range (R)	$r = \frac{R}{\bar{X}}$	1 day immersed, 28 day moist cured, unconfined compressive strength values, Individual strengths, psi				Average strength, psi	$\bar{X}=(X_2+X_3)/2$	Range (R)	$r = \frac{R}{\bar{X}}$
5	0		0	491	491	471	465	480	481	20	.0416	622	593	586	520	580	589	6	.0102
8	0			777	711	691	678	714	701	20	.0285	1043	1007	994	941	996	1000	13	.0130
11	0			1162	1109	1053	1017	1085	1081	56	.0518	1533	1421	1388	1291	1408	1424	33	.0232
5	3	1	0	517	418	474	471	486	477	7	.0147	622	606	583	583	599	595	23	.0387
	6			550	527	527	481	521	527	0	.0000	678	649	625	640	623	637	24	.0371
	9			551	546	520	520	535	533	26	.0488	698	685	685	636	676	685	0	.0000
8	3			876	833	809	790	827	821	24	.0292	1132	981	968	955	1009	974	13	.0733
	6			724	691	678	672	691	684	13	.0190	1122	971	830	777	925	900	141	.1567
	9			846	846	820	810	830	833	26	.0312	1195	1179	1109	1060	1111	1144	70	.0612
11	3			1162	1147	1045	1019	1093	1046	102	.0975	1579	1438	1395	1389	1450	1407	43	.0306
	6			1284	1175	1171	1099	1182	1173	4	.0034	1681	1622	1606	1527	1609	1614	16	.0099
	9			1244	1241	1135	1122	1186	1188	106	.0892	1632	1625	1559	1464	1570	1592	66	.0415
5	3	3	0	530	481	468	428	477	474	13	.0274	718	636	619	553	631	627	17	.0271
	6			543	514	481	435	493	497	33	.0644	721	629	590	553	642	610	39	.0639
	9			461	399	386	359	401	392	13	.0332					623			
8	3			853	688	685	659	721	667	3	.0045	1017	994	978	876	966	986	16	.0162
	6			915	872	856	853	874	864	16	.0185	1362	1277	1251	1201	1273	1264	26	.0206
	9			767	714	704	685	718	709	10	.0141	1181	1145	1079	1032	1110	1112	66	.0594
11	3			1329	1191	1145	1050	1179	1168	46	.0394	1727	1681	1606	1500	1629	1644	75	.0456
	6			1301	1291	1109	1109	1202	1200	182	.1517	2013	1842	1681	1569	1779	1762	161	.0914
	9			1261	1247	1247	1234	1247	1248	0	.0000	2085	2049	2043	1944	2130	2046	6	.0029
5	3	4	0	573	540	514	484	529	527	26	.0493	747	744	737	724	738	740	7	.0095
	6			445	438	432	428	436	435	6	.0138	708	698	665	652	681	681	33	.0485
	9			386	382	353	353	368	368	29	.0788	675	645	639	616	644	642	6	.0093
8	3			757	741	682	622	700	712	59	.0829	1095	1070	1026	991	1045	1048	44	.0420

	6			1301	1291	1109	1109	1202	1200	182	.1517	2013	1842	1681	1569	1779	1762	161	.0914
	9			1261	1247	1247	1234	1247	1248	0	.0000	2085	2049	2043	1944	2130	2046	6	.0029
5	3	4	0	573	540	514	484	529	527	26	.0493	747	744	737	724	738	740	7	.0095
	6			445	438	432	428	436	435	6	.0138	708	698	665	652	681	681	33	.0485
	9			386	382	353	353	368	368	29	.0788	675	645	639	616	644	642	6	.0093
8	3			757	741	682	622	700	712	59	.0829	1095	1070	1026	991	1045	1048	44	.0420
	6			655	652	639	599	636	645	13	.0202	1043	1043	1001	964	1013	1023	20	.0411
	9			573	547	530	514	541	538	17	.0316	941	862	846	800	862	854	87	.0187
11	3			1165	1132	1086	1066	1112	1109	46	.0415	1704	1671	1671	1556	1651	1671	33	.0000
	6			980	967	932	922	950	749	35	.0467	1497	1481	1405	1382	1441	1443	54	.0527
	9			908	899	885	774	867	892	14	.0157	1464	1329	1240	1181	1304	1285	179	.0693
5	0		0.5	688	682	649	560	645	665	33	.0496	945	928	865	803	885	897	48	.0702
8	0			855	774	747	685	765	760	27	.0355	1125	1086	1032	1010	1064	1059	66	.0510
11	0			1348	1317	1291	1093	1262	1304	26	.0199	1908	1898	1885	1737	1857	1891	17	.0059
5	3			711	691	659	632	668	675	32	.0474	1010	964	964	885	956	964	0	.0000
	6			662	645	630	590	631	638	15	.0235	1036	1022	981	948	997	1001	42	.0420
	9			662	649	639	619	642	644	10	.0155	1010	1007	987	951	989	997	20	.0201
5	3	3	0.5	780	727	701	688	724	714	16	.0224	1162	1116	1060	1032	1092	1088	56	.0515
	6			764	718	685	639	701	701	33	.0471	1251	1247	1198	1139	1209	1223	49	.0401
	9			705	685	652	632	668	668	33	.0494	1304	1291	1218	1122	1234	1254	73	.0582
5	3	4	0.5	619	603	596	593	603	599	7	.0117	912	892	872	770	862	882	20	.0227
	6			474	458	438	425	449	448	20	.0046	724	711	705	668	702	708	6	.0085
	9			609	590	567	530	574	579	23	.0397	1060	1040	1019	885	1001	1029	21	.1701

Table 7b. Tabulation of data used in determining the coefficient of variation (CV) for the study involving quadruples of strength determinations - loess soil

Cement content, %	Fly ash content, %		Sodium carbonate content, %	1 day immersed, 7 day moist cured, unconfined compressive strength values,								1 day immersed, 28 day moist cured, unconfined compressive strength values,							
	%	Lab. no.		Individual strengths, psi				Average strength, psi	$\bar{X}=(X_2+X_3)/2$	Range (R)	$r = \frac{R}{\bar{X}}$	Individual strengths, psi				Average strength, psi	$\bar{X}=(X_2+X_3)/2$	Range (R)	$r = \frac{R}{\bar{X}}$
5	0		0	257	244	228	228	239	234	16	.0684	402	382	369	381	375	375	13	.0880
8	0			422	399	386	379	396	392	3	.0077	636	622	622	629	627	622	0	.0113
11	0			573	560	517	501	538	538	43	.0799	885	882	882	869	880	882	0	.0181
5	3		0	257	251	244	238	248	247	7	.0283	399	396	386	373	388	391	10	.0665

5	0		0	257	244	228	228	239	234	16	.0684	402	382	369	381	375	375	13	.0880
8	0			422	399	386	379	396	392	3	.0077	636	622	622	629	627	622	0	.0113
11	0			573	560	517	501	538	538	43	.0799	885	882	882	869	880	882	0	.0181
5	3		0	257	251	244	238	248	247	7	.0283	399	396	386	373	388	391	10	.0665
	6			234	238	228	218	229	231	6	.0260	386	373	373	336	367	373	0	.1340
	9			238	248	238	218	235	243	10	.0412	389	363	359	353	366	361	4	.0997
8	3			359	359	353	336	352	356	6	.0169	652	613	609	606	620	611	4	.0753
	6			402	376	373	340	372	374	3	.0080	645	596	576	573	598	586	20	.0341
	9			402	376	366	363	377	371	10	.0270	672	668	606	590	634	637	62	.0973
11	3			550	550	514	511	531	532	36	.0677	952	922	885	853	903	903	35	.0388
	6			511	491	484	471	489	487	7	.0144	889	876	865	859	872	871	11	.0126
	9			444	484	484	478	485	484	0	.0000	813	754	754	751	768	754	0	.0000
5	3	3	0	261	251	241	228	245	246	10	.0407	415	409	392	392	402	400	17	.0425
	6			267	267	267	264	266	267	0	.0000					432			
	9			307	303	287	284	295	295	16	.0542	507	507	495	465	493	501	12	.0240
8	3			389	366	356	336	362	361	10	.0277	636	636	586	520	594	614	50	.0814
	6			405	389	359	356	375	374	30	.0802	688	682	678	672	680	680	4	.0059
	9			415	405	402	399	405	403	3	.0074	777	744	734	708	741	739	10	.0135
11	3			524	471	458	445	474	464	13	.0280	744	724	724	714	821	724	0	.0000
	6			606	599	592	590	599	595	7	.0118	974	945	935	908	941	940	10	.0106
	9			590	570	567	537	566	568	3	.0053	994	968	951	882	949	959	17	.0177
5	3	4	0	280	271	271	262	271	271	0	.0000	382	382	366	356	322	374	16	.0428
	6			221	218	215	182	209	216	3	.0139	373	366	366	323	357	366	0	.0000
	9			202	197	192	188	195	194	5	.0258	369	369	363	359	365	366	6	.0164
8	3			369	366	363	336	359	364	3	.0082	576	576	557	553	566	567	19	.0335
	6			343	336	326	323	332	331	10	.0302	550	520	504	501	519	512	16	.0313
	9			297	297	297	277	292	297	0	.0000	520	517	497	471	502	507	20	.0394
11	3			560	557	544	527	547	550	13	.0236	823	813	810	787	811	808	3	.0037
	6			517	504	491	488	500	497	13	.0262	826	813	797	780	804	805	16	.0199
	9			481	474	465	455	469	470	9	.0191	793	780	763	731	769	776	27	.0348
5	0		0.5	363	349	349	313	344	349	0	.0000	468	461	435	425	447	448	26	.0580
8	0			520	514	448	442	481	481	66	.1372	721	714	708	688	708	711	6	.0084
11	0			685	652	645	636	654	648	7	.0108	955	955	948	853	927	951	7	.0074

Table 7b (Continued)

Cement content, %	Fly ash content, %	Fly ash content, Lab. no.	Sodium carbonate content, %	1 day immersed, 7 day moist cured, unconfined compressive strength values, Individual strengths, psi				Average strength, psi	$\bar{X}=(X_2-X_3)/2$
5	3	4	0.5	359	346	343	343	348	345
	6			353	340	336	326	339	338
	9			323	320	310	307	315	315
5	3	1	0.5	359	349	343	320	343	346
	6			389	379	376	369	378	377
	9			379	379	363	349	368	371
5	3	3	0.5	382	373	363	356	368	368
	6			415	392	382	379	392	387
	9			455	445	425	392	429	435

st cured, gth values,			1 day immersed, 28 day moist cured, unconfined compressive strength values,							
$(X_2+X_3)/2$	Range (R)	$r = \frac{R}{\bar{X}}$	Individual strengths, psi				Average strength, psi	$\bar{X}=(X_2+X_3)/2$	Range (R)	$r = \frac{R}{\bar{X}}$
345	3	.0087	534	491	465	435	481	478	26	.0544
338	4	.0118	520	488	478	471	489	483	10	.0207
315	10	.0317	543	534	511	484	518	523	23	.0440
346	6	.0173	465	448	445	445	451	447	3	.0067
377	3	.0080	566	543	484	481	519	514	59	.1148
371	16	.0431	534	514	501	497	511	507	13	.0256
368	10	.0272	530	494	491	481	499	492	3	.0061
387	10	.0258	636	626	590	560	603	603	36	.0597
435	20	.0460	659	659	652	593	640	655	7	.0107

APPENDIX C

8a. Tabulation of data used in the study illustrating the use of control specimens in detecting outliers - Colfax soil

ıt ıt,	Fly ash content,		Molding moisture content,		7 day moist cured, 1 day immersed, individual unconfined compressive strengths, psi			Average strength, psi ($\bar{X}$)	Range (R)
	%	lab. no.	%	type^a					
	0		10.7	A	932	833	826	864	106
			12.2	A	461	432	409	434	52
			11.2	A	511	488	465	488	46
			9.9	B	1145	1142	1063	1117	82
			9.2	B	1050	1017	961	1009	89
	0		10.9	A	567	514	514	531	53
			9.9	B	846	820	816	827	30
			11.2	A	425	419	320	388	105
			10.3	B	761	711	701	724	60
			9.6	B	754	728	721	734	33
	0		11 2	A	340	313	284	312	56
			10.2	B	491	485	465	480	26
			9.8	B	491	474	445	470	46
			10.7	A	459	428	359	415	100
			11.7	A	317	238	228	259	89
	3	3	11.1	A	626	570	488	561	138
			10.0	B	770	737	649	719	121
			9.4	B	731	724	701	719	30
			11.2	A	405	329	320	368	85
			10.9	A	761	744	698	734	63
	6		11.0	A	606	557	537	566	69
			10.5	B	797	793	695	762	102
			10.0	B	728	665	603	665	125
			9.5	B	757	751	691	733	66
			11.6	A	494	415	409	439	85
	9		11.5	A	468	468	468	468	0
			10.8	A	737	682	678	699	59
			10.5	B	803	767	685	752	118
			10.0	B	754	751	701	735	53
			9.3	B	770	728	665	721	105
	3	3	11.2	A	547	533	527	536	20
			10.7	A	908	905	767	860	141
			10 2	A	1165	1093	1043	1100	122
			9.9	B	1237	1142	1135	1171	102
			9.4	B	1165	1152	1147	1155	18

		10.5	B	803	767	685	752	118
		10.0	B	754	751	701	735	53
		9.3	B	770	728	665	721	105
3	3	11.2	A	547	533	527	536	20
		10.7	A	908	905	767	860	141
		10 2	A	1165	1093	1043	1100	122
		9.9	B	1237	1142	1135	1171	102
		9.4	B	1165	1152	1147	1155	18
6		11.3	A	668	613	471	584	197
		10.9	A	961	918	·777	885	184
		10.4	B	1139	1073	987	1066	152
		9.9	B	1185	1116	1086	1129	99
		9.4	B	1157	1047	922	1042	235
9		11.4	A	895	816	793	835	102
		10.8	A	1145	1078	994	1072	151
		10.3	B	1152	1145	1073	1123	79
		9.9	B	1102	1050	1043	1065	59
		9.0	B	1143	1010	928	994	215
3	3	11.6	A	280	257	238	258	42
		11.1	A	346	343	327	339	19
		10.6	A	435	415	399	416	36
		10.2	B	507	474	465	482	42
		9.0	B	481	438	419	446	62
6		11.5	A	317	294	277	296	40
		11.0	A	376	369	354	368	17
		10.6	A	474	461	458	465	16
		9.9	A	471	468	451	463	20
		9.0	B	432	427	419	426	13
9		11.5	A	290	267	251	269	39
		11.3	A	396	356	330	360	66
		10.5	A	494	465	445	468	49
		10.2	A	520	491	474	495	46
		9.0	B	478	465	442	461	36
9	1	11.5	A	349	290	284	308	65
		11.1	A	396	356	274	342	122
		10.8	B	520	484	478	494	42
		10.1	B	573	566	563	568	10
		9.0	B	547	524	491	520	56
6		11.4	A	284	264	208	252	76
		11.2	A	310	248	229	262	81
		10.5	A	419	380	369	389	50
		10.1	B	543	524	511	526	32

		11.1	A	390	350	274	342	122
		10.8	B	520	484	478	494	42
		10.1	B	573	566	563	568	10
		9.0	B	547	524	491	520	56
6		11.4	A	284	264	208	252	76
		11.2	A	310	248	229	262	81
		10.5	A	419	380	369	389	50
		10.1	B	543	524	511	526	32
		8.7	B	540	501	445	495	95
3		11.7	A	215	208	165	196	50
		11.1	A	300	244	238	261	62
		10.4	A	408	399	379	395	29
		10.3	B	481	451	432	455	49
		9.0	B	481	465	438	461	43
3	1	11.5	A	343	320	313	325	0
		10.6	A	626	622	613	620	13
		10.2	B	675	660	649	661	26
		9.7	B	869	816	790	825	79
		8.8	B	767	764	734	755	33
6		10.6	A	737	728	682	716	55
		10.2	B	846	839	836	840	10
		9.5	B	875	780	662	773	213
		8.6	B	691	685	649	675	42
9		10.9	B	793	694	652	713	141
		10.5	B	797	780	737	771	60
		10.0	B	849	843	724	805	125
		9.4	B	807	800	790	799	17
		11.4	A	567	484	461	504	106
9	1	11.2	A	1119	1050	968	1045	151
		10.8	A	1122	1116	1093	1100	29
		10.4	B	1216	1132	1089	1146	127
		9.9	B	1119	1070	1043	1077	76

A indicates that molding moisture content was at or below optimum for maximum density.

B indicates that molding moisture content was above optimum for maximum density.

Table 8a (Continued)

Cement content, %	Fly ash content, %	Fly ash content, lab. no.	Molding moisture content, %	Molding moisture content, type	7 day moist cured, 1 day immersed, individual unconfined compressive strengths, psi			Average strength, psi ($\bar{X}$)	Range (R)	$r = \frac{R}{\bar{X}}$
11	9	1	9.2	B	1026	941	912	960	114	.1188
11	6		11.4	A	728	682	606	672	122	.1815
			11.1	A	1109	957	945	1004	104	.1633
			10.7	B	1271	1218	1129	1206	142	.1177
			9.8	B	1191	1171	1108	1156	83	.0778
			9.2	B	987	941	820	916	167	.1823
11	3		11.4	A	724	652	626	667	98	.1469
			11.0	A	1040	843	803	895	237	.2648
			10.7	A	1211	1168	1043	1141	168	.1472
			10.3	B	1297	1257	1199	1251	98	.0783
			9.1	B	1089	1077	978	1025	111	.1083
11	9	4	11.7	B	807	695	675	692	33	.0477
			12.7	B	872	737	698	764	174	.2263
			13.8	A	872	856	836	855	36	.0421
			15.4	A	691	639	550	627	141	.2249
			16.2	A	402	379	349	377	53	.1406
11	6		16.4	A	363	313	283	320	80	.2500
			14.7	A	458	445	428	444	30	.0676
			13.8	A	751	701	590	680	161	.2368
			13.1	A	961	957	925	948	36	.0380
			11.9	B	925	866	853	881	72	.0817
11	3		11.2	B	1053	1007	981	1014	72	.0710
			11.3	B	987	938	925	950	62	.0653
			12.2	A	691	662	642	665	49	.0737
			11.4	B	1073	1056	961	1030	112	.1087
			12.3	A	889	856	744	830	145	.1747
8	9	4	11.8	B	553	537	511	534	42	.0787

			11.4	B	1073	1056	961	1030	112	.1087
			12.3	A	889	856	744	830	145	.1747
8	9	4	11.8	B	553	537	511	534	42	.0787
			13.1	B	619	593	520	577	99	.1716
			13.8	B	609	599	580	596	29	.0487
			14.7	A	543	527	507	526	36	.0684
			15.6	A	267	257	244	256	23	.0898
8	6		12.0	B	685	675	652	671	33	.0492
			13.0	A	629	616	590	611	39	.0638
			13.8	A	438	369	333	380	105	.2763
			12.8	B	698	672	662	678	36	.0531
			13.5	A	576	514	514	535	62	.1159
8	3		11.0	B	711	688	668	689	43	.0624
			11.7	A	655	642	603	633	52	.0821
			12.5	A	507	491	488	495	19	.0384
			11.2	B	734	724	724	728	10	.0137
			12.8	A	434	425	422	427	13	.0304
5	9	4	12.0	B	373	356	323	351	50	.1425
			13.3	B	399	389	382	390	17	.0436
			14.0	A	386	376	376	379	10	.0264
			14.5	A	340	336	326	334	14	.0419
			15.3	A	290	277	264	277	26	.0939
5	6		11.4	B	419	412	399	410	20	.0488
			12.4	B	432	415	409	418	23	.0550
			12.9	A	412	402	389	401	23	.0574
			13.5	A	336	310	290	312	46	.1474
			14.3	A	234	205	188	209	46	.2201
5	3		10.8	B	511	507	504	507	7	.0138
			11.8	A	464	451	448	455	16	.0352
			12.0	A	435	432	419	428	16	.0374
			13.2	A	257	254	195	235	62	.2638
			10.4	B	458	438	432	443	26	.0588
8	3	3	10.9	B	682	665	659	668	23	.0344
			11.1	A	682	645	639	655	43	.0656

Table 8b. Tabulation of data used in the study illustrating the use of control specimens in detecting outliers - loess soil

Cement content, %	Fly ash content,		Molding moisture content,[a]		7 day moist cured, 1 day immersed, individual unconfined compressive strengths, psi			Average strength, psi ($\bar{X}$)	Range (R)	$r = \frac{R}{\bar{X}}$
	%	lab. no.	%	type						
5	0		13.7	B	247	241	221	237	26	.1097
			14.6	B	254	238	231	241	23	.0954
			16.3	A	261	248	241	250	20	.0800
			18.7	A	109	100	96	102	13	.1275
			17.7	A	169	169	169	169	0	.0000
8	0		14.9	B	389	373	330	364	69	.1621
			16.2	B	428	389	369	396	69	.1490
			17.2	A	349	336	287	324	62	.1914
			17.2	A	257	257	251	255	6	.0235
11	0		16.8	A	392	392	386	390	6	.0154
			16.4	B	533	511	488	511	45	.0881
			16.8	A	524	505	445	491	79	.1069
			17.6	A	300	287	267	285	33	.1158
5	3	1	16.0	B	277	274	267	273	10	.0366
			16.7	B	271	271	265	269	6	.0223
			17.2	A	231	228	221	227	10	.0441
			18.0	A	162	159	161	131	3	.0186
			18.4	A	132	123	109	121	23	.1901
5	9		15.2	B	303	300	284	296	19	.0642
			16.4	B	336	336	333	335	3	.0090
			17.0	A	300	287	274	287	26	.0906

---

[a] A indicates that molding moisture content was at or below optimum for maximum density.

B indicates that molding moisture content was above optimum for maximum density.

Table 8b (Continued)

Cement content, %	Fly ash content, %	Fly ash content, lab. no.	Molding moisture content, %	Molding moisture content, type	7 day moist cured, 1 day immersed, individual unconfined compressive strengths, psi			Average strength, psi ($\bar{X}$)	Range (R)	$r = \frac{R}{\bar{X}}$
5	9	1	17.6	A	238	231	175	215	63	.2930
			16.0	B	300	287	280	289	20	.0692
11	9		16.0	B	513	504	481	500	32	.0640
			15.9	B	511	514	465	493	46	.0933
			16.9	B	550	504	379	478	171	.3577
			17.2	B	458	454	415	443	43	.0971
			18.0	A	343	317	248	302	95	.3146
			19.4	A	202	182	172	185	30	.1622
11	6		16.0	B	619	573	563	585	56	.0957
			16.6	B	570	563	557	563	13	.0231
			17.6	A	422	392	317	377	105	.2785
			19.4	A	202	182	172	185	30	.1622
11	3		14.9	B	543	524	445	504	98	.1994
			15.7	B	576	553	491	540	85	.1574
			16.9	A	514	438	438	463	76	.1641
			17.5	A	363	363	300	342	63	.1842
			17.8	A	297	294	261	284	36	.1268
8	6		15.3	B	412	402	392	402	20	.0498
			16.0	B	409	405	396	403	13	.0323
			16.9	A	380	349	343	357	37	.1036
			18.0	A	247	231	165	215	82	.3814
			18.3	A	211	195	195	200	16	.0800
			15.0	B	537	531	517	528	20	.0379
8	3		15.3	B	369	349	343	354	26	.0734
			15.9	B	415	392	363	390	52	.1333
			17.0	A	389	382	330	377	59	.0565
			17.9	A	264	254	215	244	49	.2008
			18.6	A	192	179	172	181	20	.1105
8	9		14.5	B	392	386	382	387	10	.0258

								390	52	.1333
			17.0	A	389	382	330	377	59	.0565
			17.9	A	264	254	215	244	49	.2008
			18.6	A	192	179	172	181	20	.1105
8	9		14.5	B	392	386	382	387	10	.0258
			15.7	B	399	392	297	363	102	.2810
			16.5	B	392	389	379	387	13	.0336
			17.2	A	369	353	343	355	26	.0732
			18.2	A	211	202	198	204	13	.0637
5	9	4	15.2	B	218	205	202	208	16	.0769
			16.1	B	205	202	198	202	7	.0347
			16.9	B	211	211	208	210	3	.0143
			18.1	B	225	218	211	218	14	.0642
			21.2	A	215	208	198	207	17	.0821
5	6		15.7	B	225	225	221	223	4	.0179
			16.8	B	241	241	228	237	13	.0549
			17.7	B	238	225	221	228	17	.0746
			18.8	B	248	238	231	239	17	.0711
			20.5	A	179	159	155	164	24	.1463
5	3		16.9	B	241	231	228	233	13	.0558
			18.3	B	241	238	238	239	3	.0126
			19.1	A	169	165	155	163	14	.0859
			20.3	A	100	100	96	98	4	.0408
			17.7	B	218	218	211	216	7	.0324
11	3		16.0	B	465	455	438	452	27	.0597
			16.6	B	478	474	458	470	20	.0426
			17.8	A	534	507	497	513	37	.0721
			18.5	A	429	419	412	420	17	.0405
			19.3	A	287	287	271	282	16	.0567
11	9		17.0	B	419	405	386	403	33	.0818
			17.8	B	379	373	366	372	13	.0349
			20.0	B	402	402	402	402	0	.0000
			21.4	A	409	399	379	395	30	.0759
			18.8	B	399	389	346	378	53	.1402
11	6		16.5	B	425	409	409	414	16	.0386
			17.9	B	435	425	422	427	13	.0304
			19.0	B	445	438	435	439	10	.0228
			19.9	A	438	435	419	431	19	.0441
			20.0	A	294	240	257	280	37	.1321

			19.9	A	438	435	419	431	19	.0441
			20.0	A	294	240	257	280	37	.1321
8	3		16.1	B	359	356	349	355	10	.0282
			16.7	B	376	369	334	361	37	.1025
			17.5	B	346	343	340	343	6	.0175
			19.7	A	175	165	149	163	26	.1595
			18.5	A	277	274	274	275	3	.0109
8	9		17.2	B	313	294	287	298	26	.0872
			18.0	B	300	294	280	291	20	.0687
			19.5	A	290	280	277	283	13	.0459
			20.0	A	303	303	290	299	13	.0435
			18.5	B	290	287	200	259	90	.3475
8	6		17.1	B	317	313	307	312	10	.0321
			18.0	B	343	333	313	330	30	.0909
			19.2	A	317	317	317	317	0	.0000
			19.3	A	343	333	323	333	20	.0601
8	9	3	16.1	B	353	323	317	331	36	.1088
			16.8	B	363	353	333	344	30	.0860
			17.9	A	290	287	284	287	6	.0204
			18.8	A	172	152	149	158	23	.1456
			17.3	B	365	349	349	355	16	.0451
8	6		16.2	B	354	343	333	345	26	.0754
			16.4	B	386	386	354	377	27	.0716
			17.3	A	382	363	363	369	19	.0515
			18.3	A	221	215	162	199	59	.2965
			17.9	A	277	257	244	260	33	.1269
8	3		14.8	B	362	359	344	357	13	.0364
			16.0	B	379	364	353	367	26	.0708
			16.7	B	366	363	363	364	3	.0082
			18.1	A	228	215	198	214	30	.1402
			18.6	A	175	159	142	159	33	.2075
5	3	3	15.1	B	228	215	220	220	13	.0591
			15.7	B	257	254	248	253	9	.0356
			16.9	B	254	251	244	250	10	.0400
			18.0	A	175	175	165	172	10	.0581
			17.5	A	237	221	211	223	26	.1166
5	9		15.1	B	234	231	218	228	16	.0702
			16.2	B	241	234	231	235	10	.0426
			16.7	B	251	251	231	244	20	.0820

Table 8b (Continued)

Cement content, %	Fly ash content,		Molding moisture content,		7 day moist cured immersed, indiv unconfined compr strengths, psi	
	%	lab. no.	%	type		
5	9	3	17.3	A	231	229
			17.8	A	221	221
5	6		14.9	B	241	238
			15.9	B	244	231
			16.5	B	248	248
			17.6	A	238	234
			17.8	A	228	221
11	9		14.7	B	497	432
			15.8	B	501	497
			18.0	A	359	300
			18.6	A	231	218
11	6		15.9	B	514	495
			16.7	B	494	471
			17.6	A	421	376
			18.9	A	274	244
			22.1	A	553	543
11	3		15.5	B	497	474
			16.5	B	534	530
			17.3	A	471	428
			18.5	A	208	204
			17.0	A	533	517

y moist cured, 1 day mersed, individual onfined compressive strengths, psi		Average strength, psi ($\bar{X}$)	Range (R)	$r = \frac{R}{\bar{X}}$
229	221	227	10	.0441
221	204	216	17	.0787
238	231	237	10	.0422
231	231	235	13	.0553
248	241	245	7	.0286
234	215	229	23	.1004
221	188	212	40	.1887
432	428	452	69	.1527
497	455	484	46	.0950
300	280	313	79	.2524
218	205	218	26	.1193
495	484	497	30	.0604
471	479	479	23	.0480
376	369	389	52	.1337
244	238	252	36	.1429
543	494	530	59	.1113
474	442	471	55	.1168
530	504	523	30	.0574
428	280	393	191	.4860
204	198	204	10	.0490
517	514	522	19	.0364