HIGHWAY QUALITY CONTROL PROGRAM

MICHIGAN
HIGHWAY RESEARCH LABORATORY SECTION

MICHIGAN DEPARTMENT OF STATE HIGHWAYS

HIGHWAY QUALITY CONTROL PROGRAM

C. D. Church
F. Copple
C. A. Zapata

Final Report on a Highway Planning and Research Investigation
Conducted in Cooperation with the U. S. Department of Transportation
Federal Highway Administration

Research Laboratory Section
Testing and Research Division
Research Project 63 G-123
Research Report No. R-779

Michigan State Highway Commission
Charles H. Hewitt, Chairman; Louis A. Fisher, Vice-Chairman
Claude J. Tobin; E. V. Erickson; Henrik E. Stafseth, Director
Lansing, September 1971

CONTENTS

Page

SUMMARY

This study was undertaken to upgrade current Michigan highway standard specifications and acceptance testing procedures where indicated, by determining realistic tolerances based on statistical quality assurance concepts. An examination of historical test results for certain quality characteristics and limited field experimental studies were to form the bases for any recommendations.

The introduction briefly outlines the purpose, objectives and accomplishments of the study.

A new sampling and testing procedure for concrete flexural strength based on modulus-of-rupture is presented. Daily testing of shorter beams is recommended. Based on an analysis of historical strength test data, a sample acceptance level (SAL) is determined. A field study, undertaken to test the new procedures, indicated that they were workable and would provide better estimates of concrete strength.

Preformed neoprene joint sealant is the next item investigated. Results of past acceptance testing are summarized. Alternative sampling approaches are discussed and each quality characteristic is classified according to seriousness should it not meet specifications. Suggestions for an improved quality assurance function are given.

A proposed procedure for acceptance testing of Class III Granular Material is presented. Operating characteristic curves, sample size, warning procedures, check tests, sample acceptance limits, etc., are described.

A proposal for a 22A aggregate gradation field experiment is included even though the study was not carried out. Size and location of experimental projects, controlled variables, type of tests, and sampling procedure are all considered.

Phase I of the project was described in a prior report, which is included as Appendix A. Here, specific characteristics of highway materials and construction for which statistical quality assurance methods would be practical and advantageous are delineated. From historical records, general operating levels and variability of these characteristics were determined. Results provided workable ideas for field experiments under actual job conditions. Historical records are summarized in tabular and graphical form. Results of a limited field experiment on aggregate gradation are included. Recommendations primarily involved procedural rather than statistical considerations. Another field study was conducted on the uniformity of transit-mix concrete. Several parameters were investigated. Performance of concretes containing different additives is compared.

I
INTRODUCTION

Highway departments exist to provide the public with a system of pavements and bridges constructed and maintained as economically as possible; adequate to carry vehicles from point to point. To carry out this function, some form of quality control or quality assurance has always been, and must continue to be, used by these departments.

The first decisions necessary to the control or assurance of quality are those concerned with highway design. After design, there is the major problem of ensuring conformance of the product to the design specifications; and that problem is the subject of this study. Even though products have been manufactured or constructed to the same specifications, their degree of conformance may be very different. The capability of a process determines largely how a product conforms to specifications. For example, if a concrete mixing process were capable of producing concrete whose compressive strength could be controlled only within ± 500 psi, it couldn't possibly conform well to a specification permitting a tolerance of ± 300 psi. Thus, one major concern of quality control is to determine the capabilities of processes to meet specifications. Process capability studies are normally a problem for the producer, while the consumer is interested primarily in assuring that the goods he accepts meet his specifications. However, since highway departments (the consumers) may pay wastefully higher prices if their specifications are too restrictive, it is important that specifications be practical and realistic with respect to process capabilities.

Traditionally, specifications did not spell out the actual needs of a design nor reflect the capabilities of the construction process or variation of materials. Quality requirements in such specifications were often based on judgement and experience. It is apparent that good highways were, and continue to be, built using traditional specifications as applied by experienced engineers and inspectors, capable of recognizing poor materials and workmanship. Under traditional procedures, one "representative" sample is taken periodically and the test result is taken as a valid indication of the quality of the material being inspected. If the test result falls within specification limits, the material is accepted. If it falls outside specified limits, the engineer or inspector decides whether to test again, reject the product, or accept it on the basis of substantial compliance.

Practical specifications should be clear, unambiguous, and designed to provide the desired end product at a minimum cost. To be realistic, specifications must be written to allow for the inevitable variations in materials and construction or manufacturing processes. Legally, it is the responsibility of the originator to make the meaning of each specification clear and incapable of misrepresentation. It is often impossible, however, to foresee every contingency; so some judgements must be left to the discretion of the engineer. But for administrative efficiency, and to minimize both misunderstandings in the field and the additions of contingency items to a contractor's bid, specifications in all possible cases should be practical, realistic, and specific.

This study was initiated to develop a quality assurance program that would minimize the need for intuitive decision making, provide more uniformity in acceptance testing procedures, and generate practical and realistic specifications in precise legal or contractual terms. It officially began on July 1, 1963 as a two-year Highway Planning and Research project in cooperation with the Federal Highway Administration. With FHWA approval, the project was extended to June 30, 1969. The first two years are called Phase I and subsequent work is said to be Phase II. The purpose and objectives of the project as stated in the proposal are as follows:

Purpose

> "The general purpose of the project is to upgrade current MDSH standard specifications where indicated, by the determination and adoption of realistic tolerances based on statistical concepts, and to develop guidelines to be followed in the development of future specifications.
>
> The study will be designed and conducted so that as the work progressed and significant results were obtained, they would be presented immediately as recommendations for consideration in connection with appropriate specification changes."

Objectives

> "1. To delineate specific areas pertaining to highway materials and construction where the application of quality control concepts will be practical and advantageous.
>
> 2. To develop suitable acceptance plans in these areas.
>
> 3. To rewrite existing materials and construction specifications in accordance with approved acceptance plans.

4. To prepare a Department Manual containing guidelines for continuing improvements in sampling, testing, and acceptance in relation to highway construction, maintenance, and operations and to reduce variability of performance among contractors conducting similar work. "

Research Procedure

Briefly, the research procedure as outlined in the proposal is:

"1. Review of the overall acceptance testing program for highway materials and construction as required by current Michigan specifications. Also review the decision-making principles presently practiced by the Department.

2. Establish promising areas for special study and research from historical data and from recent sources. Determine where and when improved quality or better job control is needed.

3. Develop practical and meaningful specification limits for improved acceptance sampling.

4. Conduct statistically designed laboratory or field experiments to determine specification limits where proper operating limits are not known.

5. Start the development and operation of a Sigma Bank (or standard deviation bank) to provide values of Sigma (standard deviation) that might be expected under given field conditions. Sigma values, together with related material and construction data, will be stored on IBM cards for rapid recovery in connection with future specification writing.

A continuing effort will be made to secure more recent construction, laboratory, and field test results. As these test results become available, they will be analyzed and evaluated. When significant results are obtained, they will be presented as recommendations for upgrading current highway specifications. "

Accomplishments

During the first two-year span of the project, historical data were analyzed for several materials, and one field test--involving aggregate grain size analysis--was completed. Data from the analysis of historical data, and the field test published previously as "Highway Quality Control Program: Statistical Parameters" (MDSH Research Report No. R-572, March 1966) are included in this report as Appendix A. There are major weaknesses in using historical data, e.g., absence of information regarding sources or causes of variations, and the lack of a sound random sampling plan during the time when data were gathered in the field.

In early 1966, a Departmental Advisory Committee was appointed to assume general administration of the project. Familiarity with statistical quality control procedures was not a prerequisite for membership, but Committee members were selected on the basis of their expertise in a field where statistical quality control or assurance might be implemented. The Committee was composed of the following representatives:

Chairman: R. L. Greenman, Testing and Research Engineer
E. A. Finney, Director, Research Laboratory
L. T. Oehler, Asst. Director, Research Laboratory
C. J. Olsen, Director, Testing Laboratory
J. C. Brehler, Director, Field Testing
C. M. Ellis, Asst. to the Bridge Construction Engineer
D. L. Wickham, Asst. to the Road Construction Engineer
E. M. Noble, Manager, Highway Planning and Research

The priority list developed by the Committee was as follows:

1. Concrete modulus-of-rupture testing
2. Air and slump tests for concrete, with particular emphasis on slag-aggregate concrete
3. Dimensions of neoprene seals
4. Surfacing aggregates: 21A, 22B, 22C, 22D, 22E, 23A, and 24A
5. Bituminous aggregate 20A
6. Sand 2NS
7. Thickness of selected subbase (under concrete pavement)
8. Thickness of concrete pavement
9. Density tests on embankment construction
10. Porous material, Grade A (loss by washing)
11. Surface tolerance on concrete pavements
12. Penetration of recovered asphalt
13. Height of fence fabric.

This report discusses Priority Items 1 and 3, and parts of Items 4, and 10; the only ones on the list which were studied during the duration of this project. Although time did not permit study of the remaining items, they are listed in order to show where the Advisory Committee believed emphasis was needed in the area of quality control. Even though it was relatively low on the list, Item 10 was included in the study because of a special request from the Construction Division.

For reasons discussed earlier in the report, historical data were not considered satisfactory for developing realistic acceptance testing plans. Therefore, a proposal for a large-scale field test investigating grain size distribution of 22A surfacing aggregate was prepared and submitted to the Committee for approval. After reviewing the proposal, which is included in the report, a majority of the Committee declared the suggested field test redundant and it was not approved.

In the spring of 1969, about eight hours of lectures on "Statistical Quality Control" were presented by the Research Laboratory to several engineers in the Construction Division. Their questions and comments indicated that the construction engineers who attended were enthusiastic, and apparently convinced of the value of statistical quality control. However, persons attending the class were below the top management levels where decisions to use statistical methods must utltimately be made.

II

TESTING FOR CONCRETE FLEXURAL STRENGTH BASED ON MODULUS-OF-RUPTURE

(PRIORITY ITEM 1)

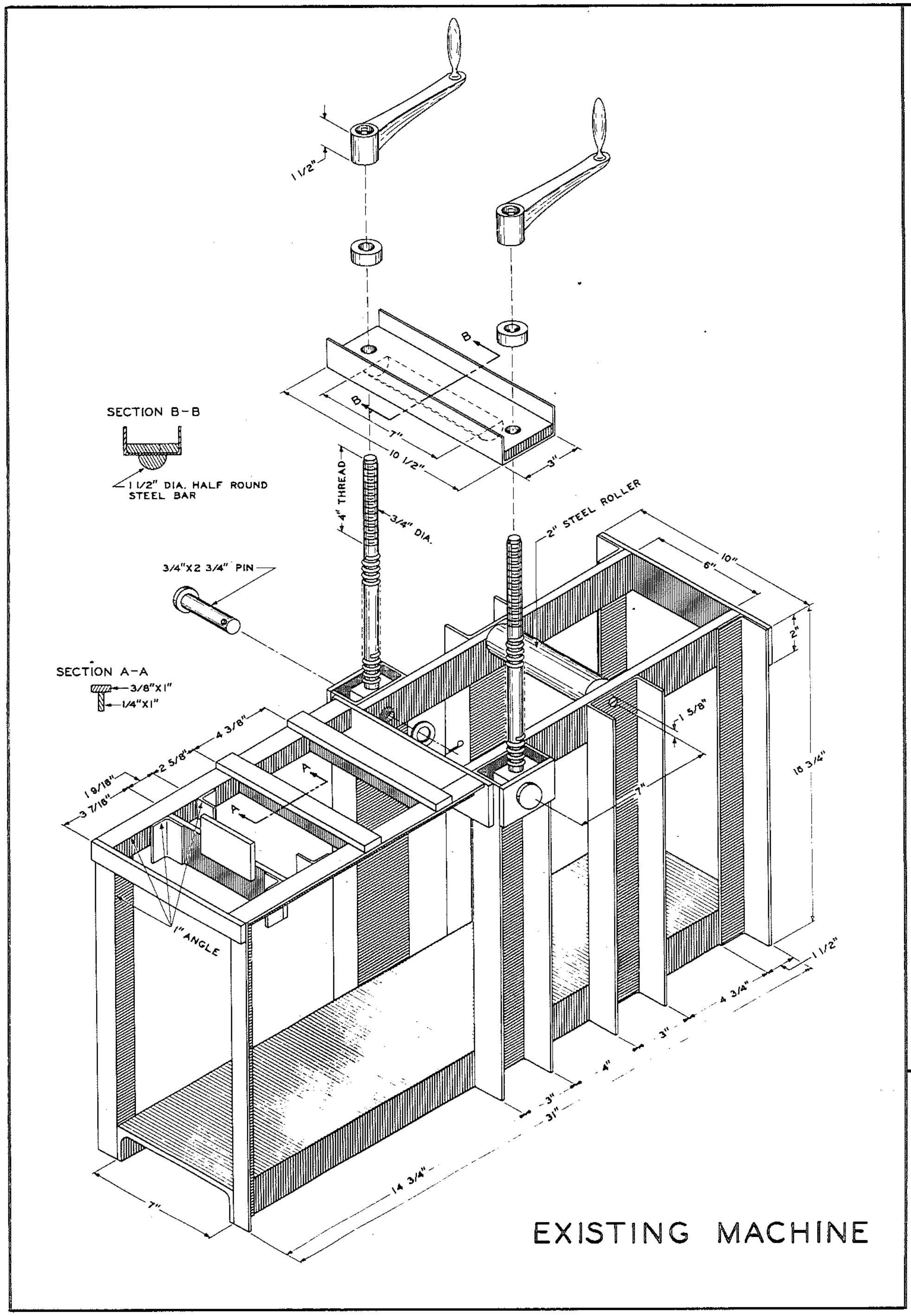
The Committee's Priority Item 1 was the application of quality assurance principles to a modulus-of-rupture test to determine flexural strength of pavement and bridge concrete. In Michigan, flexural strength tests are used to gather information concerning concrete mix design and the quality of its constituents. If further investigation indicates, for example, that low flexural strength is due to structural or chemical composition of the aggregate, this will be sufficient reason for rejecting aggregate from this source for further use.

The question immediately arose as to whether concrete flexural strength was a sufficiently critical material characteristic to warrant imposition of these new techniques. In reply to this, it can be pointed out that since concrete strength is presently being checked by the Department for decision making purposes, statistical methodology can be appropriately applied in an attempt to increase confidence in the test. The methodology can be used regardless of the purpose of the inspection, i.e., as a check on contractual conformance or, as presently used, to gather information on the mix design.

Perhaps the justification for priority consideration is that modulus-of-rupture testing is one of the least sensitive areas for study in that it would involve minimum interference with on-going construction. It was felt that experience gained in the development of a program for this characteristic of relatively minor importance would be beneficial to future plans for more critical items concerning contractual obligations.

Background

The study began with an analysis of historical modulus-of-rupture data; the results of which, along with additional data, were presented in a Progress Report (MDSH Research Report No. R-550, October 1965). In June 1967, a proposal (Research Report No. R-583) for revising modulus-of-rupture acceptance procedures was submitted. This proposal was based on additional historical within-project data analysis and embodied new sampling and statistical considerations. As a result of this proposal,



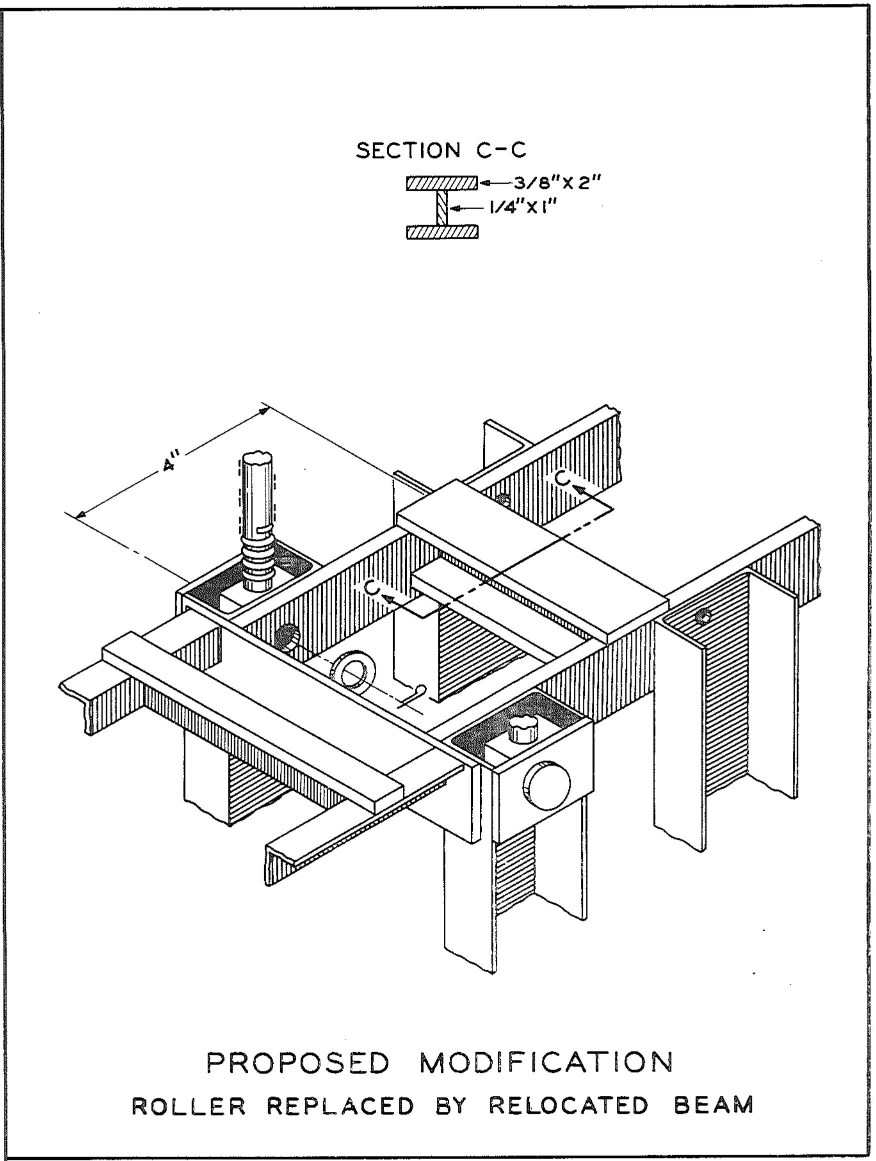


Figure 1. Concrete flexural strength testing machine.

a field study beginning in 1968 was undertaken, essentially to determine the operational feasibility of the recommended procedures. Its intent was not necessarily to indentify factors influencing concrete strength.

It should be noted that, for the purposes of the field study, consideration was limited to pavement concrete (as opposed to bridge concrete) and also to seven-day tests.

Initial Considerations

Time Lag - It is difficult to design a sound quality assurance plan for concrete strength from the standpoint of taking immediate corrective action since the tests are made at least seven days after the concrete has been placed; usually too late to be of any influence on the quality of the related mix. In addition, any confirmed strength deficiency is not easy to physically rectify. Nevertheless, we believe more meaningful quality assurance procedures can be implemented if more reliable estimates of concrete strength are desired.

Shorter Beams - Application of statistical quality assurance procedures usually involves an increase in sample size relative to current requirements. In order that this be accomplished with a minimum increase in inconvenience, it was recommended that shorter, single-break beams be molded. Moreover, experience has shown that with the heavier, longer beams used previously, a number of tests were lost because of broken beams resulting from the rough handling that goes with more difficult lifting.

Conventional 6- by 6- by 36-in. beams would be abandoned and an increased number of individual 6- by 6- by 18-in. beams be substituted, thus allowing more independent tests to be made.

Because of the shorter beam length, certain modifications would be necessary to the beam breaking frame. Minimum alteration would involve moving a support roller, further modification could make the test frame less cumbersome (Fig. 1).

The benefits of 18-in. beams outweigh the inconvenience of having to handle more beams to achieve an equal number of tests. Advantages include: easier handling, less chance of damage in handling, and a minimum loss of information if a beam should be damaged. Also, it would seem to be more psychologically appealing to base acceptance-rejection decisions on several sample units. By far the most significant benefit, however, is the improved quality assurance reliability gained, since the use of smaller beams allows a larger number to be conveniently tested. In addition, two separate beams are more desirable since two breaks of the same beam do not give the scope of information that two breaks of different beams do. This is especially true if several beams are widely and systematically selected, thereby representing different portions of the concrete mix.

Specifications - It is difficult to determine sample acceptance limits, risk levels, desired quality, etc., from an engineering point of view. Statistically based specifications require more information than the current single-valued specification calls for. We must know, for example, whether this specification represents a maximum or minimum average, or desired quality level, and what proportion below desired quality can be tolerated over the long run.

For modulus-of-rupture, the current specifications call for a seven-day strength of 550 psi. Is this an absolute minimum value, or would a 550 psi average indicate adequate strength? If a minimum, must each and every test meet or exceed this value, or is there a certain proportion of tests below this value that could be tolerated? The design of a statistically based acceptance procedure must consider these questions and should provide their answers.

Lot Size - Another difficulty in developing an acceptance and testing procedure based on modulus-of-rupture measurements is the establishment of the amount of concrete production to be represented by a sample. This problem is complicated by the nature of the production process. Since material flow is continuous, discrete lots are not naturally defined, and construction on a specific project is sometimes sporadic.

For strength testing, proper quality assurance procedures require daily sampling when construction is in progress instead of the current every other day practice. Daily samples would facilitate the detection of any major strength differences that might occur due to any day-to-day production changes. It was decided, then, to focus attention on a day's production for inspection purposes. Any further division into smaller "lots" would be purely arbitrary and, hence, unnecessary.

Operating Characteristic

Based on analyses of limited historical data, a rough initial estimate of within-project standard deviation was obtained. This estimate was used as one of the parameters defining the operating characteristic (OC) established as a model for the field study. In some cases, data showed estimates of standard deviation to range as high as 95 psi. If OC's are used to determine acceptance limits, then the standard deviation is an important factor and estimates thereof should be as reliable as possible. Therefore, continuous analysis of actual test data would be desirable in order to confirm or update current estimates.

Several plans are discussed and illustrated by OC curves in the discussion that follows. They represent various combinations of sample size for the average of two breaks per beam, and quality levels for a particular contractor's risk ($\alpha = .05$) and consumer's risk ($\beta = .10$). Other combinations of risk levels could be used if appropriate, which might affect sample size and sample acceptance limit (SAL) or both.

In practice, the procedure would be to compute the average modulus-of-rupture value of the sample and compare it with the plan's SAL. Acceptance or rejection of the lot is then based on whether or not this level was exceeded by the sample average.

It should be noted that for most of the cases described in this report the OC curves were not designed with a fixed desired quality level (DQL) in mind. A DQL is a predetermined value of the quality characteristic for which the probability of acceptance will be set high.

Historical evidence indicates that about 10 percent of the data are falling below the specification limit of 550 psi but are being tolerated. This is based on the distribution of individual breaks and also the average of two breaks per beam for accepted results. The curves in Figure 2 were designed with these facts in mind; that is, so that the chance of accepting concrete with strength of 550 psi is 10 percent (consumer's risk).

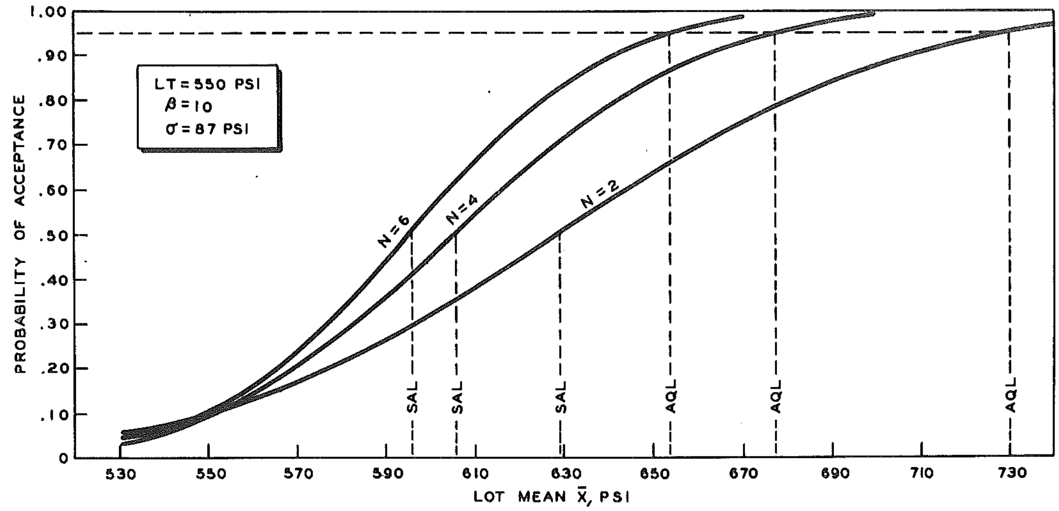


Figure 2. Operating characteristic curves for three sample sizes (N).

The curves also illustrate the different probabilities of acceptance for concrete of a given strength with appropriate shifts in the SAL as sample sizes change, but for a fixed standard deviation. Note that the various sample acceptance levels are considerably above 550 psi. This means that more rejections would occur and, thus, more laboratory analyses, etc., would be required under any of these plans than was experienced in the past if the actual strength remained at a level less than the SAL. Currently, laboratory analyses are supposed to be made on each beam whose flexural strength is below 550 psi.

Perhaps 550 psi is not the critical lot tolerance (LT) but rather is an average desired quality level. If this is the case, curves can be designed fixing the SAL at 550 psi. With the SAL fixed at 550 psi, half of the lots actually having this modulus-of-rupture value will be accepted and half will be rejected. Curves for this case and for various sample sizes are shown in Figure 3. Higher probabilities of accepting lots of less than 550 psi are quite evident. For comparison purposes, Figure 3 also shows OC curve 'A' illustrating a plan for which the DQL is fixed at 550 psi and the minimum tolerance at 470 psi. Note that the required sample size is 10.

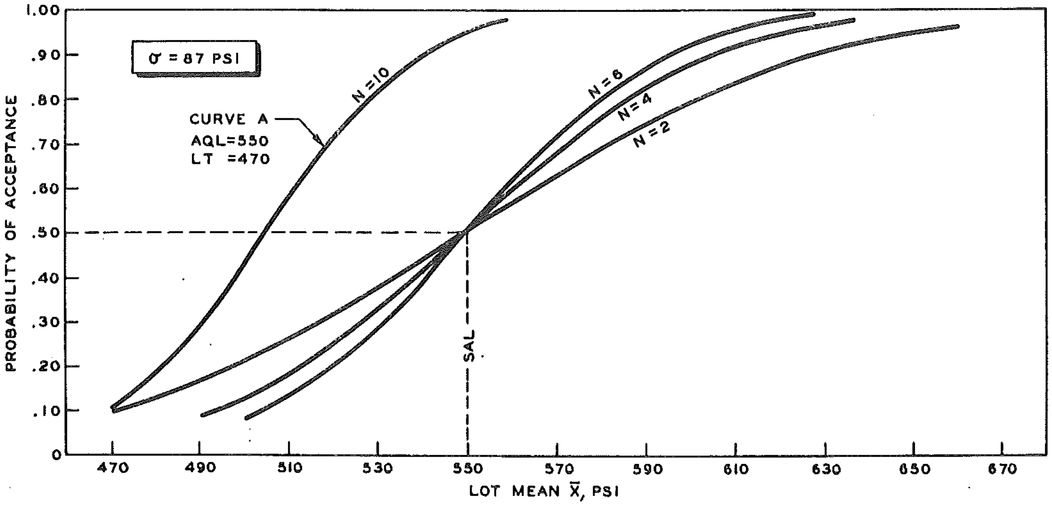


Figure 3. Operating characteristic curves for three sample sizes (N) all designed to have a SAL of 550 psi.

It has been established that there are slight seasonal differences in the variation of modulus-of-rupture distributions which would affect lot variance and thus require different sample sizes for each season. Under ideal conditions, variation should be the same from lot to lot. If the apparent difference is deemed significant, alternative plans would have to be followed each season. Assuming again that 550 psi is the LT, Figure 4 shows how the sample acceptance levels change for spring, summer, and fall. With the sample size fixed, the difference observed is caused by the seasonal variation. A plan appearing essentially the same for all three seasons could be used in which the DQL and LT are fixed. However, the sample size required for this constant degree of protection might vary by one or two units for each different season.

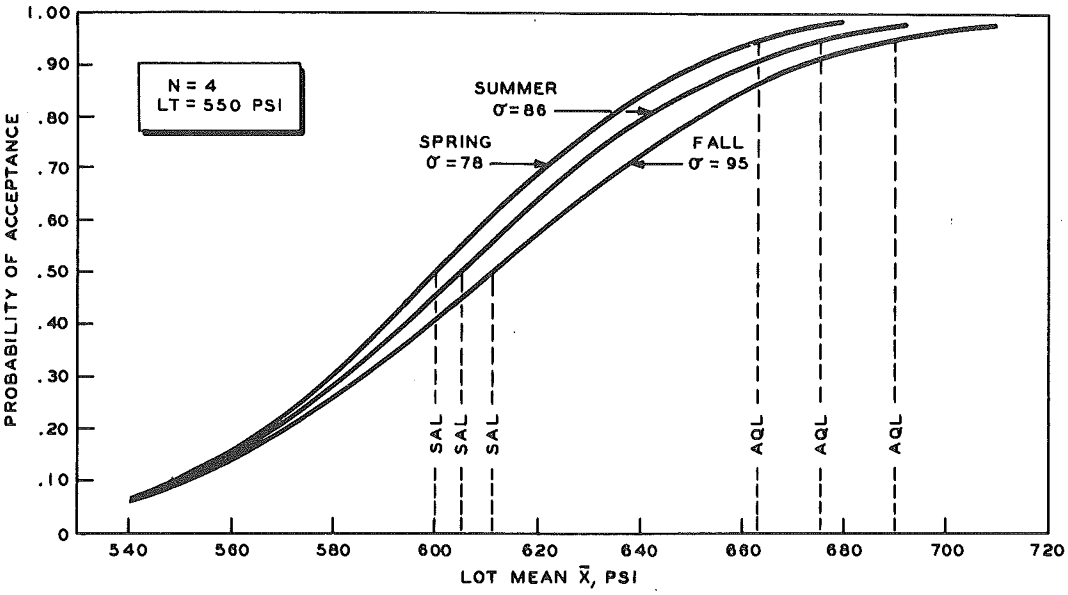


Figure 4. Operating characteristic curves based on the variability with each season.

It has been suggested that because of seasonal differences, a correction might be applied to non-typical seasonal modulus-of-rupture values. The correction would be based on curing conditions; primarily temperature. In view of the complications involved in estimating and administering such a factor, it appears more practical to use the alternative plan which would require taking perhaps one more beam or meeting a slightly higher sample average. Under either of these procedures, however, the improved power of discrimination between good and bad lots does not seem significant enough to warrant switching plans between seasons. A uniform plan for all conditions, representing a compromise between administrative difficulty and the risk incurred, appears more feasible at the present time.

FIELD STUDY

To determine the feasibility of the proposed new procedures, a field study was undertaken. This study was conducted over a two-year period (1968-69) under actual job conditions.

Operating Characteristics

The specific construction projects involved consisted of portions of I 496 in Ingham and Eaton Counties in Michigan. Strength specifications for the project were not affected by the proposed changes, only the procedures for sampling were modified.

Figure 5 shows the OC curve for the model sampling plan developed to give example parameters for this study. The curve is defined by an estimated standard deviation of modulus-of-rupture values based on past within-project variations and assumes that 550 psi is a minimum acceptable strength. This is different than estimates defining previous curves which included between-project and between-year variability.

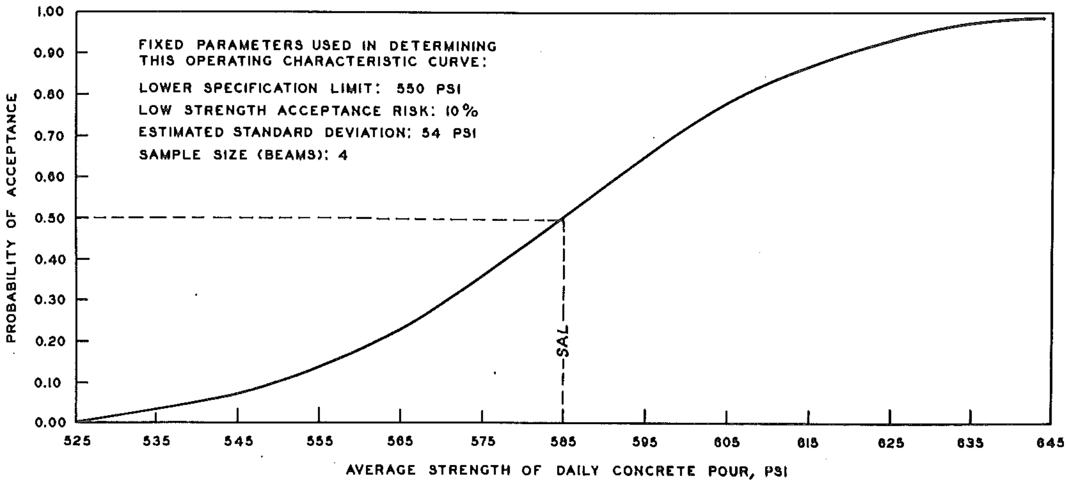


Figure 5. Operating characteristic curve for field study.

Under this plan, the average modulus-of-rupture of four beams must be 585 psi or more at the age of seven days. If not, concrete from the day's pour will be judged deficient in strength and appropriate action taken. Meanwhile, the pavement would not be opened to traffic pending verification.

The 585 psi is not a new strength specification. Rather, it is a SAL that must be met or exceeded by the sample average in order to have a specified degree of confidence that the concrete has at least the required 550 psi flexural strength. In other words, given a sample size of four beams, and assuming a standard deviation in strength measurements of 54 psi, then a minimum sample average of 585 psi should be met in order that the risk of accepting low strength (below 550 psi) pavement never exceeds 10 percent (in this example) in the long run.

For practical as well as statistical considerations it was determined that a daily sample size of four beams was satisfactory and feasible. Individual beams of the set of four were to be cast at approximately even intervals during the day, i.e., early and late morning and early and late afternoon.

Job Control Specimen

Another aspect of this study was that two 18-in. beams were formed at the same time in existing 36-in. molds. Thus, four beam pairs were obtained, resulting in eight test beams per day. The existing 36-in. molds were modified by placing a transverse partition at the center in order to provide two separate 18-in. beams.

Following routine identical initial curing (24 hr), the beams were removed from the mold and prepared for final curing. One beam was cured in the usual manner at the site by burying it in moist sand for the remaining five days. The other beam of the pair (designated a job control specimen) was transported to a location indoors expected to be more conducive to optimum curing conditions. Hypothetically, ideal temperature control would be achieved. In addition, these beams were sealed in polyethylene bags to insure moisture retention.

It was hoped that under these conditions, the job control beams would approach the maximum strength for seven days and that comparisons with field-cured beams would indicate any limitations in the field method. Also, if this procedure were feasible, higher strengths, more representative of ultimate pavement valves, should be observed, thus minimizing (if not eliminating) the need to justify to the FHWA any low-strength test results obtained using field-cured beams.

Results

Data were collected according to the new procedures during the summer construction seasons of 1968 and 1969. For descriptive purposes, the modulus-of-rupture values were plotted in chronological sequence for each year-curing method combination as illustrated by Figures 6 and 7.

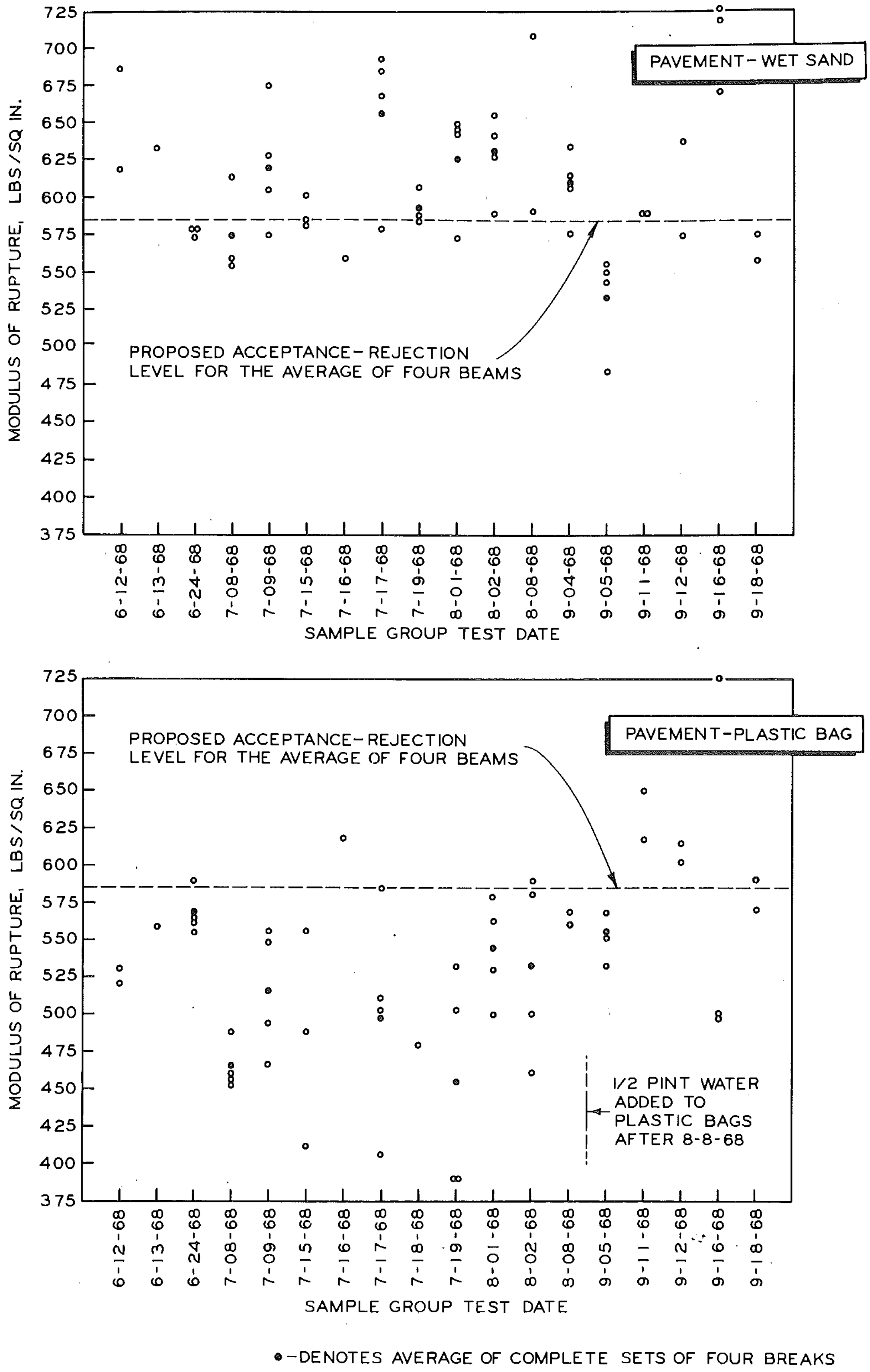


Figure 6. Chronological sequence of 1968 test results for each cure method.

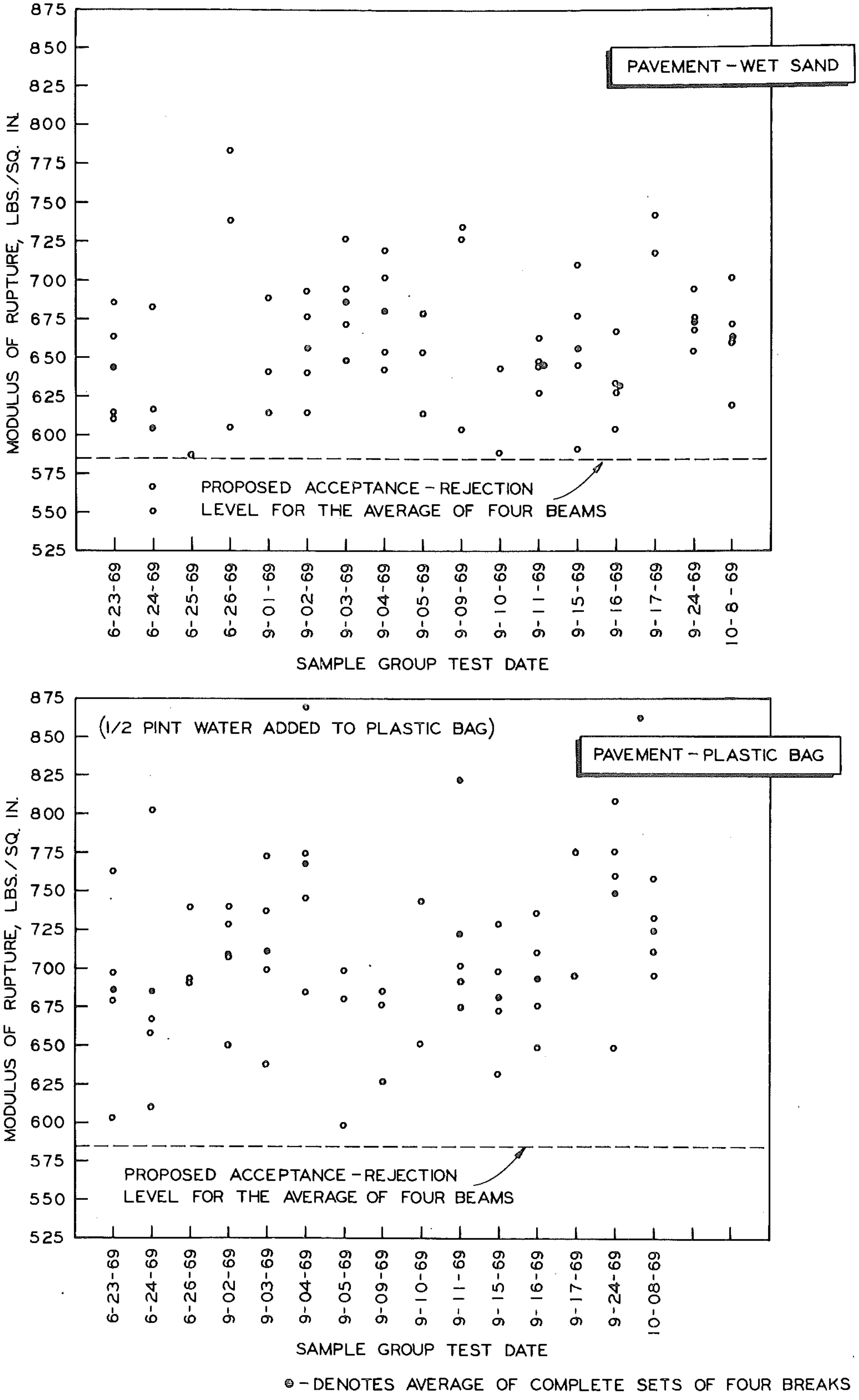


Figure 7. Chronological sequence of 1969 test results for each cure method.

The acceptance limit which appears on the charts is for reference purposes only, and does not apply to individual breaks. In practice, however, this limit would apply to the average of the four breaks per sample. For the present we can estimate from the chart whether or not the average would fall below the limit and, if so, if this was caused by a single break or by the whole sample being low.

Several important observations can immediately be made. First, by noting the dates the test samples were molded, the discontinuity of construction is apparent. This may or may not be of any practical significance in this case. In general, however, ideal quality control is achieved when the production process is continuous.

Another observation is that there is a relatively high within-sample variation. This result is disturbing since it renders more difficult the ability to detect real between-lot shifts in concrete strength, which is the main objective of acceptance sampling. It should be evident, however, how unreliable a single break would be in estimating concrete strength, and at the same time the superiority of using an average of several breaks.

How much this within-sample variation is due to testing and curing and how much represents actual strength differences cannot be determined from this study. For the 1969 data, however, 12 percent more within-sample variation was noted[1] in the strengths obtained from the beams cured in polyethylene bags than in the field-cured beams. Since the same concrete was used for each set, the implication is that the difference is due to factors other than inherent strength variation, namely, curing or testing or both.

The fact remains that the variation is present and must be taken into account when prescribing sample size. A controlled experiment could serve to isolate the various contributions to this variance. In the case of modulus-of-rupture testing it is difficult to foresee what practical measures could be taken which would decrease any assignable error. Hence, the high within-sample variation will have to be recognized and most likely will be reflected in an increased sample size requirement.

A comparison of charts for each of the year and cure method combinations reveals substantial differences among the overall average levels. There are reasonable explanations for these differences as mentioned in the following discussion.

---

[1] As measured by component sum of squared deviations from the sample means.

It will also be noted that, on some days, construction was limited and fewer than four beams were obtained. The proposed acceptance plan was designed for a sample size of four. Theoretically, each different sample size would have its own different SAL. We feel that this situation would be unworkable in the field. Therefore, we propose that a minimum of four beams be molded regardless of the amount of construction on a given day. The inspector would have to have some advance knowledge, of course, in order to plan his sampling intervals. If advance notice was not available as to when construction would be completed, his sample would have to be fulfilled from the last concrete available.

An additional descriptive illustration of the data is shown in Figure 8. Here the frequency distributions of the modulus-of-rupture values are presented for each of the sub-group combinations. From these distributions an idea of the magnitude of the variation of the data can be obtained as well as the relative position of the distributions on the modulus-of-rupture scale.

Although additional data would have been desirable to allow a more firm conclusion, comparisons between the observed standard deviations (especially for the field data) and the estimate based on past data (54 psi) which was used in the model OC curve indicate relatively close agreement.

Polyethylene Bag Curing

The scheme involving the final curing in polyethylene bags of one beam of the 18-in. pair did not work out satisfactorily. According to the initial proposal, storage and testing was to take place at the Project Engineer's office or other suitable building near the construction site. In some cases this involved considerable extra handling and transporting.

As the charts indicate for the tests made in 1968, strengths of the beams cured in this manner were, on the average, much lower than those cured as usual. These results are quite contrary to expectations. As the results became known, we felt that the reason might be due to tears in the bag or inadequate sealing, allowing the beams to dry too fast. Subsequently, a cup of water was added to each bag before sealing in order to remedy the problem, but at too late a date to improve the 1968 findings since construction was discontinued shortly thereafter.

The whole procedure of transporting the beams, placing them in polyethylene bags, and testing, appears to have turned out to be quite bothersome to field personnel and did not generate adequate concern for their care. It appears that final curing in plastic bags in the field requires too much delicate care and attention to be a practical routine field practice.

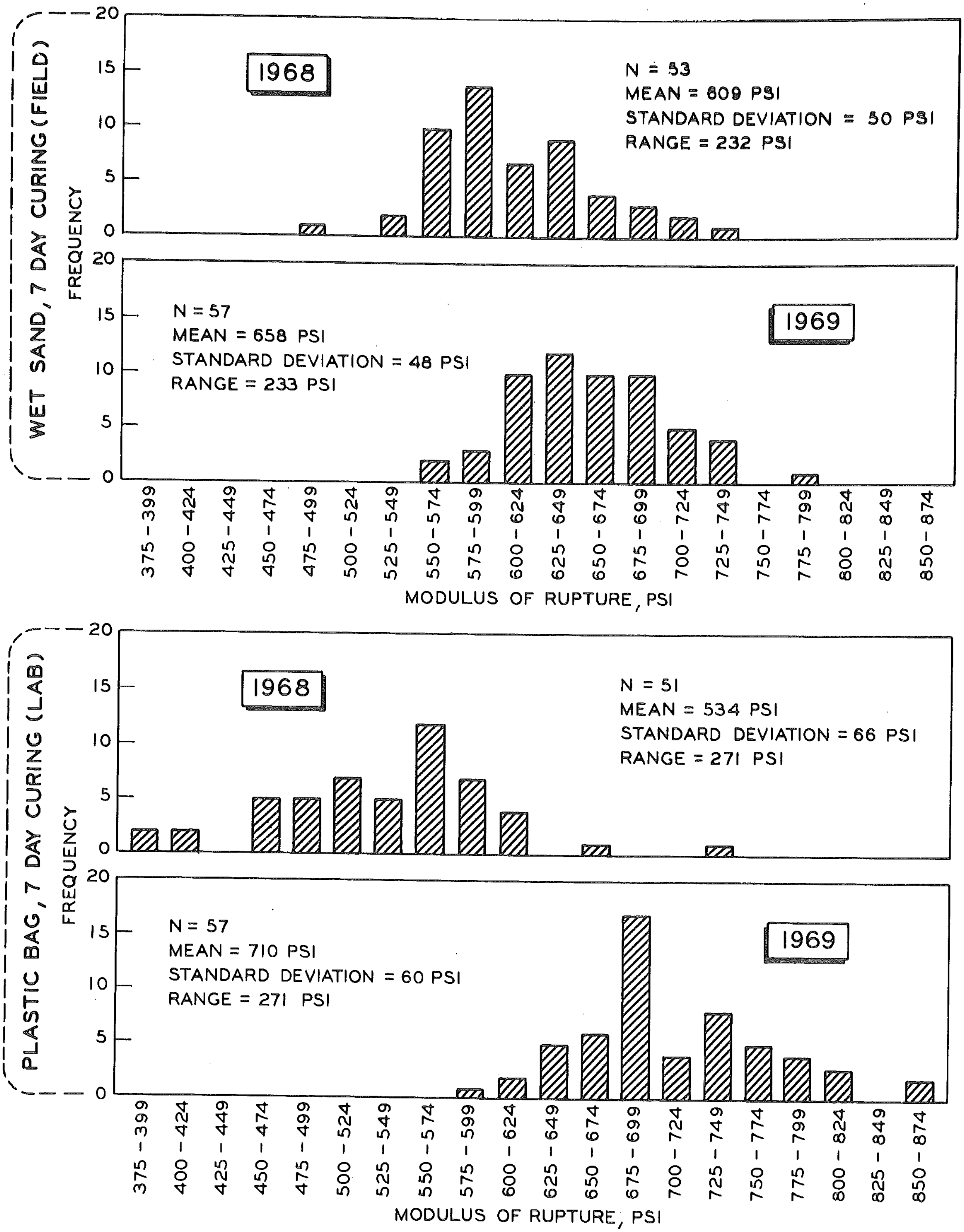


Figure 8. Comparison of test results for each cure method by year.

For the tests made in 1969, the beams designated for the plastic bags were brought to the Research Laboratory for final curing. Here extreme care was taken in handling and testing the beams. Water was routinely added to the bags before sealing. Continuous monitoring of room temperature and bag integrity were made.

When housed at the Research Laboratory, strengths were found to be high; on the average, at a higher level than comparable beams cured in the field. As a routine practice, however, final curing at the Research Laboratory is not feasible due to the transportation involved. In this study the construction project was near enough that distance was no problem.

Besides the 1968-69 difference in final curing conditions for beams sealed in polyethylene bags, concrete with low slump was used in the 1969 constructions. This was required because the slip-form paving technique was used on the project. Lower slump concrete may have resulted in a higher overall average modulus-of-rupture but would not seem to totally account for the large difference (176 psi).

Field Operation

The field operation in and of itself (excluding the plastic bag operation) seems quite feasible. Minor procedural questions arose such as how many times to rod the mix in each section of the modified 36-in. mold, and the difficulty in removing the 18-in. beams from the mold due to the separating partition. However, operational procedures such as these would be completely specified when and if this procedure is adopted. For the time being, each of the 18-in. sections was rodded half as much as the usual for the 36-in. molds. Also, the use of 18-in. molds in lieu of modified 36-in. molds will eliminate the removal difficulty.

CONCLUSIONS

The main conclusion drawn from this limited trial is that the proposed field sampling and testing procedures are workable. Moreover, this plan would provide better estimates of concrete strength due to increased daily sample sizes.

Final curing of a duplicate set of beams in sealed polyethylene bags does not appear practical. The delicate handling required, the transportation involved, and inadequate or unavailable space in field offices--all combined--tend to rule against this practice. Of course, if the value of this information is great enough, extreme and costly procedures could be imposed to accomplish the desired handling and curing conditions.

Adopting a meaningful quality control program requires more than an improved sampling scheme. Continuous surveillance and analysis of test results should also be conducted. Initial estimates of process parameters may have to be adjusted, desired quality levels may change, or other circumstances might be altered, any of which could necessitate procedural changes. Hence, if and when statistical quality assurance plans are implemented, provisions for administrative follow-up studies should also be established at the same time for a really effective quality control program.

III
PREFORMED JOINT SEALANTS

(PRIORITY ITEM 3)

Second in order of priority, as established by the Steering Committee, was concrete air content and slump testing. However, historical data were not suitable for statistical analysis so, for expediency, it was decided to investigate the third item of priority, preformed neoprene joint seals.

Current specifications call for preformed sealer of virgin, crystallization resistant neoprene which is flexible, pliable, and will retain elasticity and other physical properties under temperature and various other constraint conditions. It must pass several tests ranging from tensile strength to ozone resistance and, in addition, meet dimensional requirements (Table 1).

Designing sampling plans for inspection of neoprene sealer to assure conformance, presents interesting problems in the economics of measurement, common to most statistical acceptance plans. The problems arise from considerations concerning classification of the sampling plan as an attributes or variables plan and the selection of a random sample from a lot.

The distinction between variables and attributes sampling plans can be described briefly as follows. When a record is made of an actual measured quality characteristic, such as a dimension expressed in thousandths of an inch, the quality is said to be expressed by variables. When a record shows only the number of articles conforming or failing to conform to a specified requirement, it is said to be a record by attributes. This applies, for example, to many things that may be judged only by visual examination, or by "go" and "no go" gages. In addition, however, some characteristics that are specified as measurable variables are inspected merely as conforming or non-conforming to specifications.

The two systems compare roughly as follows:

Characteristic	Attribute	Variable
Number of observations for good data	many	few
Information value per observation	low	high
Overall cost per observation	low	high
Speed of use	fast	slow

TABLE 1
ABRIDGED 1970 MICHIGAN SPECIFICATIONS

1-1/4 in. CONTRACTION SEAL	
Overall width	1-1/4 in. min
Operational flat area of each side of seal when compressed to 1 in.	7/8 in. max
Overall depth of seal when compressed to 1/2 in.	2 in. max
1-5/8 in. EXPANSION SEAL	
Overall width	1-5/8 in. min
Operation flat area of each side of seal in uncompressed condition	1 in. min
Overall depth of seal when compressed to 1-1/2 in.	1-3/8 in. min
Overall depth of seal when compressed to 5/8 in.	2 in. max

Tensile strength, psi	2,000 min	ASTM D412 (die C or D)
Elongation at break, percent	250 min	ASTM D412
Hardness, Type A durometer	60 ± 5	ASTM D2240[1]
Oven aging, 70 hr @ 212 F		
Tensile strength, change, percent	20 max	
Elongation, change, percent	20 max	
Hardness, points change	0 to 10	
Oil swell, 70 hr @ 212 F, ASTM oil 3		ASTM D471
Weight change, percent	45 max	
Ozone resistance, 70 hr @ 104 F		ASTM D1149
20 percent strain, 300 pphm in air	No cracks	
Compression recovery, 50 percent deflection, percent of original width		Departmental methods
70 hr @ 212 F	85 min[2]	
70 hr @ 14 F	88 min[2]	
22 hr @ -20 F	83 min[2]	
Compressive force vs. deflection[3]		Departmental methods
1-1/4 in. Contraction Seal		
compressed in 1 in., lb per lin in.	4.0 min	
compressed in 1-1/2 in., lb per lin in.	35.0 max	
1-5/8 in. Expansion Seal		
compressed to 1-1/2 in., lb per lin in.	4.0 min	
compressed to 5/8 in., lb per lin in.	25.0 max	

---

[1] Specimens for the hardness tests shall consist of strips cut from the webs or walls of the seal. If the specimen is not at least 0.12 in. thick, two or more plies shall be used to obtain this thickness.

[2] Adhesion between any of the webs or cracking of any of the webs shall mean the sample has failed the compression recovery test.

[3] Compressive force vs. deflection tests made only on transverse seals for concrete pavement.

The important advantage of a variables sampling inspection plan over an attribute sampling inspection plan is that for any desired degree of protection or risk, fewer items have to be inspected to judge the acceptability of a lot. If inspection of the item is costly, requires a great deal of inspection labor and expensive equipment, or if the item is damaged or destroyed by the test, use of a variables plan will be found profitable. Testing neoprene joint seal fits all the prerequisites in favor of a variables inspection plan.

The extent of conformance or non-conformance to the desired value of a quality characteristic is given weight where variables criteria are used. This may be important where there is a margin of safety in the design specifications, or a zone of indifference between clearly satisfactory and clearly unsatisfactory product. Finally, variables information usually gives a better basis for guidance toward quality improvement--provided by control charts for averages.

Perhaps the most serious limitation on the use of variables plans, however, is that acceptance criteria must be applied separately to each quality characteristic. And, as mentioned, there are at least 14 quality characteristics that neoprene joint seal has to meet. If all these characteristics are to be examined at a given inspection, a single set of attributes sampling criteria could be applied in the acceptance decision. By contrast, if each characteristic is subject to variables inspection, 14 different sets of criteria must be employed.

In practice, however, it is common to group attributes into an allowable number of defects of critical major and minor categories and to sample accordingly. Unless trivial defects are to be given the same weight as serious ones, it is essential to have a classification of defects or a classification of characteristics. Hence, here too we have multiple criteria to be applied.

It should be mentioned that not all of the requirements for neoprene joint seal would seem to be equally critical. Those most critical, bearing directly on in-service performance, are perhaps compression-recovery characteristics and physical dimensions. Others, such as tensile strength, might be tested for information only, since joint seal normally undergoes no appreciable tension in service and joint seal with insufficient strength is seldom encountered.

This being the case, the test may merely lend additional assurance regarding the general durability of the material. (Low tensile strength, for example, may indicate an inadequate cure.) For this class of requirements, reduced inspection is warranted, possibly on a spot-check basis.

It will be noted in Tables 3 and 4 that only a small percentage of material was rejected on the basis of these criteria.

Classifying Quality Characteristics for Seriousness

In order to determine if any difference exists in the importance of the various quality characteristics of neoprene joint seal, two experts were asked to classify them according to the following definitions:[2]

Class I - Very Serious

a) Will surely cause an operating failure of the seal in service which cannot be corrected in the field.
b) Will render seal totally unfit for service.

Class II - Serious

a) Will probably cause operating failure of the seal in service which cannot be readily corrected in the field.
b) Will surely involve increased maintenance or decreased life.
c) Will cause increase in installation effort.

Class III - Moderately Serious

a) May possibly cause an operating failure of the seal in service.
b) Likely to cause trouble of less serious nature than operating failure, such as substandard performance.
c) Likely to involve increased maintenance or decreased life.
d) Will cause minor increase in installation effort.

Class IV - Not Serious

a) Will not affect operation or performance, maintenance, life of the seal in service.
b) Will not affect installation effort.
c) Minor defects of appearance, finish, or workmanship.

The results of this classification are given in Table 2.

While, admittedly, two judges are a small sample upon which to base conclusions, some observations can be made which, if generally held,

---

[2] After J. M. Juran, Quality Control Handbook, McGraw-Hill, New York, 1962, Sec. 8.

TABLE 2
RATINGS OF IMPORTANCE OF CHARACTERISTICS OF NEOPRENE JOINT SEAL

Quality Characteristic		Classification*					
		Longitudinal		Transverse Contr. & Constr.		Transverse Expansion	
		C_1	C_2	C_1	C_2	C_1	C_2
Width	Over	II	II	II	II	III	II
	Under	II	III	II	III	II	III
Depth	Over	II	IV	II	II	II	III
	Under	III	IV	III	IV	III	III
Tensile strength		III	III	III	III	III	III
Elongation at break		III	III	III	III	III	III
Permanent set at break		III	II	III	II	III	II
Recovery:							
70 hr at 212 F		I	III	I	III	I	III
70 hr at 14 F		I	II	I	II	I	II
22 hr at -20 F		I	II	I	II	I	II
Heat aged:							
Tensile strength		II	III	II	III	II	III
Elongation		II	III	II	III	II	III
Hardness		II	III	II	III	II	III
Oil swell, 70 hr at 212 F		II	III	II	III	II	III
Ozone resistance		II	II	II	II	II	II

* C_1 and C_2 represent two different judges.

could have a bearing on acceptance plans. One is that it makes no difference to these judges what the intended use of the seal is (longitudinal, transverse, etc.); a specific characteristic is judged equally serious for each type of seal.

Judge C_1 classified each characteristic almost consistently one level more serious than Judge C_2. Also, with one exception, all characteristics were rated at least "moderately serious." Finally, one judge classified only three characteristics as "very serious."

The classification results seem to indicate that all characteristics are somewhat equally critical. No clear-cut, consistent sub-grouping is apparent. This suggests that to establish non-compliance with the specification for the neoprene joint seal as a whole, it should be both necessary and sufficient to establish non-compliance for any single quality characteristic.

Sampling Plans

Since tests made to determine the acceptability of the neoprene joint sealer for most of the characteristics involve measurement on a continuous scale, it seems unwise not to take advantage of this information by simply recording passing or failing. Moreover, there would be no economic gain as far as the test procedures are concerned if an attributes rather than variables plan were adopted. On the contrary, given the same test methods and procedures, more tests would be necessary under an attributes plan, thus increasing overall cost. Ozone resistance is the only characteristic that would be considered an attribute since the specifications call for the appearance of "no cracks."

The main task to be accomplished using a variables plan is that of designing statistically sound acceptance criteria for each of the quality characteristics. But this is a preliminary consideration. The final administration of the variables plan should not be particularly more difficult than would an attributes plan. Increased difficulty would arise in designing plans if inspection was to be on an attributes basis. The very low percent defective of most of the characteristics combined with small lot and sample size, would necessitate assuming non-normal distributions which would complicate computation of acceptance-rejection limits.

The formation of lots (i.e., groups of reels) and the random selection of testing samples from the individual, randomly selected reels, present further difficulties. The irregular way in which the sealer is presented for inspection complicates adoption of a statistically sound acceptance sampling plan. Presentation is "irregular" in that lots vary in size and variable lengths are on the reels.

If the Department were willing to revise their existing sampling methods, a procedure could be adopted that would allow a more genuinely random selection of test samples from the reels of neoprene seal. Under present practices, the manufacturer could conceivably manipulate the quality by more strictly controlling the production of the seals near each end of the reel, where he knows the samples will be taken. We cannot, however, indiscriminately or purely at random, select samples, as this would result in useless short lengths of seal. Perhaps steps could be taken to insure a uniform length to be contained on each reel such that samples of required lengths could be systematically taken at intervals that would allow virtually complete use of the remaining pieces, despite the fact that our pavement widths are not uniform.

Comparing Table 3 with Table 4 reveals that much improvement has been achieved in the overall quality of neoprene joint seal since 1965. A very high percentage now receives unconditional approval and outright rejections have been reduced significantly. Expansion joint seal has a poorer record than contraction joint seal as far as unconditional approvals go. This appears to be primarily due to minor deficiencies in elongation properties. Rejections for this type of seal were low, however.

Regarding dimension-related measurements, here, as well, few problems were encountered. No rejections were made and only a few measurements did not meet with unconditional approval. New dimensional specifications have evolved and are in effect which are essentially minimum requirements, i.e., 1.250-in. minimum width as opposed to the previous 1.250 ± 0.062 in. width. Additional depth and operational flat area requirements must be met when the test specimen is under compression. The remaining physical requirements remain substantially unchanged in the current (1970) specifications from the 1967 requirements.

Conclusions

In order to improve the effectiveness and/or efficiency of the quality assurance function for preformed neoprene joint seal, certain changes in procedure could be adopted. Whether or not these changes are warranted in view of performance history of the seal in service is not examined here.

The changes might include:

1) Establishment of varying sampling rates consistent with the seriousness of the quality characteristic.

2) Some lots are recommended for use even with one or more minor deficiencies judged to be harmless. It should be determined what effect any interaction of the deficiencies might be, and a limit established as to the number and type allowed if this practice is to continue.

3) Base dimension acceptance criteria on averages and statistically determined limits rather than an arbitrary number of passing measurements out of an arbitrary total number of measurements as in the current practice.

TABLE 3
SUMMARY OF JOINT SEAL ACCEPTANCE TEST RESULTS FOR 1965

	1-1/4 in. Contraction Material						1-5/8 in. Expansion Material			
	Manufacturer A		Manufacturer B		Manufacturer C		Manufacturer D		Manufacturer E	
	Length ft	Percent of total	Length ft	Percent of total	Length ft	Percent of total	Length ft	Percent of total	Length ft	Percent of total
Approved	1,497	2.1	---	---	---	---	371	4.2	---	---
Recommended*	50,800	70.0	4,278	80.9	4,965	100.0	6,142	70.0	---	---
Rejected	20,219	27.9	1,010	19.1	---	---	2,264	25.8	800	100.0
Total	72,519	100.0	5,288	100.0	4,965	100.0	8,777	100.0	800	100.0
Physical dimensions	23,511	32.4	250	5.8	---	---	1,095	12.5	---	---
Hardness, physical dimensions	8,962	12.4	---	---	---	---	---	---	---	---
Hardness	738	1.0	---	---	---	---	---	---	---	---
Hardness, elongation	238	0.3	---	---	---	---	---	---	---	---
Hardness, elongation depth	732	1.0	---	---	---	---	---	---	---	---
Hardness, width	10,000	13.8	---	---	---	---	---	---	---	---
Elongation, depth	575	0.8	---	---	---	---	---	---	---	---
Depth	1,145	1.6	---	---	---	---	597	6.8	---	---
Width	1,510	2.1	---	---	3,750	75.5	4,270	48.7	---	---
Width, depth	---	---	---	---	---	---	180	2.0	---	---
Width, elongation	825	1.1	---	---	---	---	---	---	---	---
Recovery	269	0.4	---	---	---	---	---	---	---	---
Recovery, width, depth	2,295	3.2	---	---	---	---	---	---	---	---
Recovery, tensile strength depth	---	---	4,028	76.0	---	---	---	---	---	---
Recovery, depth	---	---	---	---	1,215	24.5	---	---	---	---
Physical dimensions	1,682	2.3	---	---	---	---	---	---	---	---
Compression, elongation	---	---	1,010	19.1	---	---	---	---	---	---
Compression, recovery	---	---	---	---	---	---	1,164	13.3	800	100.0
Depth	15,478	21.4	---	---	---	---	---	---	---	---
Width	1,703	2.4	---	---	---	---	203	2.3	---	---
Longitudinal cracks	1,356	1.9	---	---	---	---	897	10.2	---	---

* "Recommended" sealant had deficiencies minor enough for conditional approval, or met specifications as revised after the contract was let.

TABLE 4
SUMMARY OF JOINT SEAL ACCEPTANCE TEST RESULTS FOR 1968-1969

	1-1/4 in. Contraction Material						1-5/8 in. Expansion Material					
	Manufacturer A		Manufacturer B		Manufacturer C		Manufacturer A		Manufacturer B		Manufacturer C	
	Length ft	Percent of total	Length ft	Percent of total	Length ft	Percent of total	Length ft	Percent of total	Length ft	Percent of total	Length ft	Percent of total
Approved	193,799	90.4	108,117	82.6	55,165	93.3	33,916	70.8	16,228	74.9	7,665	62.5
Recommended*	7,790	3.6	18,629	14.2	575	1.0	13,959	29.2	5,439	25.1	3,050	24.8
Rejected	13,075	6.0	4,220	3.2	3,390	5.7	0	0	0	0	1,550	12.7
Total	214,664		130,966		59,130		47,875		21,667		12,265	
Elongation, depth	---	---	---	---	---	---	11,509	24.4	3,909	18.0	---	---
Depth	---	---	3,743	2.9	75	0.1	---	---	1,215	5.6	---	---
Width	---	---	14,886	11.3	---	---	---	---	---	---	---	---
Width, depth	4,290	2.0	---	---	---	---	---	---	---	---	---	---
Recovery, width, depth	---	---	---	---	---	---	---	---	315	1.5	---	---
Recovery, depth	---	---	---	---	---	---	2,450	4.8	---	---	---	---
Elongation	3,500	1.6	---	---	500	0.9	---	---	---	---	3,050	24.8
Compression, elongation	---	---	---	---	890	1.5	---	---	---	---	1,550	12.7
Compression	13,075	6.0	---	---	2,500	4.2	---	---	---	---	---	---
Exposure to Ozone	---	---	4,220	3.2	---	---	---	---	---	---	---	---

* "Recommended" sealant had deficiencies minor enough for conditional approval, or met specifications as revised after the contract was let.

4) Examine the possible overlap of certain tests. It may be that a certain test is in effect duplicating the information gained in another, although different parameters are measured.

5) Consider changing the current practice of sampling at one (or both) ends of the reel by devising a sampling procedure that enables samples to be removed from inner parts of the reel without undue waste.

IV

ACCEPTANCE TESTING OF CLASS III GRANULAR MATERIAL

(PRIORITY ITEM 10)

On April 22, 1968 the Research Laboratory was asked to develop a statistical sampling plan for job control of Class III Granular Material. Because of its variability, the material--being used for bridge abutment backfill--was not well suited for conventional "representative" sampling, and field problems were being experienced as to when to decide to reject a quantity. Therefore, although it was ranked tenth on the priority list, attention was immediately given to developing a statistical acceptance testing plan for the loss by washing requirement of Class III Granular Material.

The Department's Standard Specifications (1967) provided grading requirements based on U. S. Standard Sieves as follows:

Total Percent Passing 3-in. Sieve - 100
Total Percent Loss by Washing - 0 to 10

This latter to be determined on that portion of the sample which passes the 1-1/2-in. sieve, in accordance with the "Method of Test for Amount of Material Finer than No. 200 Sieve in Aggregate," AASHO T-11.

On June 6, 1968 a statistical acceptance testing plan was submitted to the Testing and Research Engineer. This plan, a test for loss by washing as originally presented is included.

Since measurement values are taken on a continuous scale to evaluate the desired properties (gradation), a variables acceptance plan was recommended. Statistical parameters were developed from data taken during job control tests in early 1968. Material tested came from four different sources used during construction of I 496 in or near Lansing. Frequency distributions (Fig. 9) show the material to be quite variable so an acceptance plan based upon individual test values would be virtually meaningless. Values of total material passing the No. 200 sieve include weights for both loss by washing and a subsequent "dry shake." It is believed that, given an adequate supply of wash-water in the field, the material passing the dry shake would have been washed through. The current upper limit of 10 percent loss by washing was taken to be the desired quality level where chances of rejecting a lot which did not exceed that proportion would be very small.

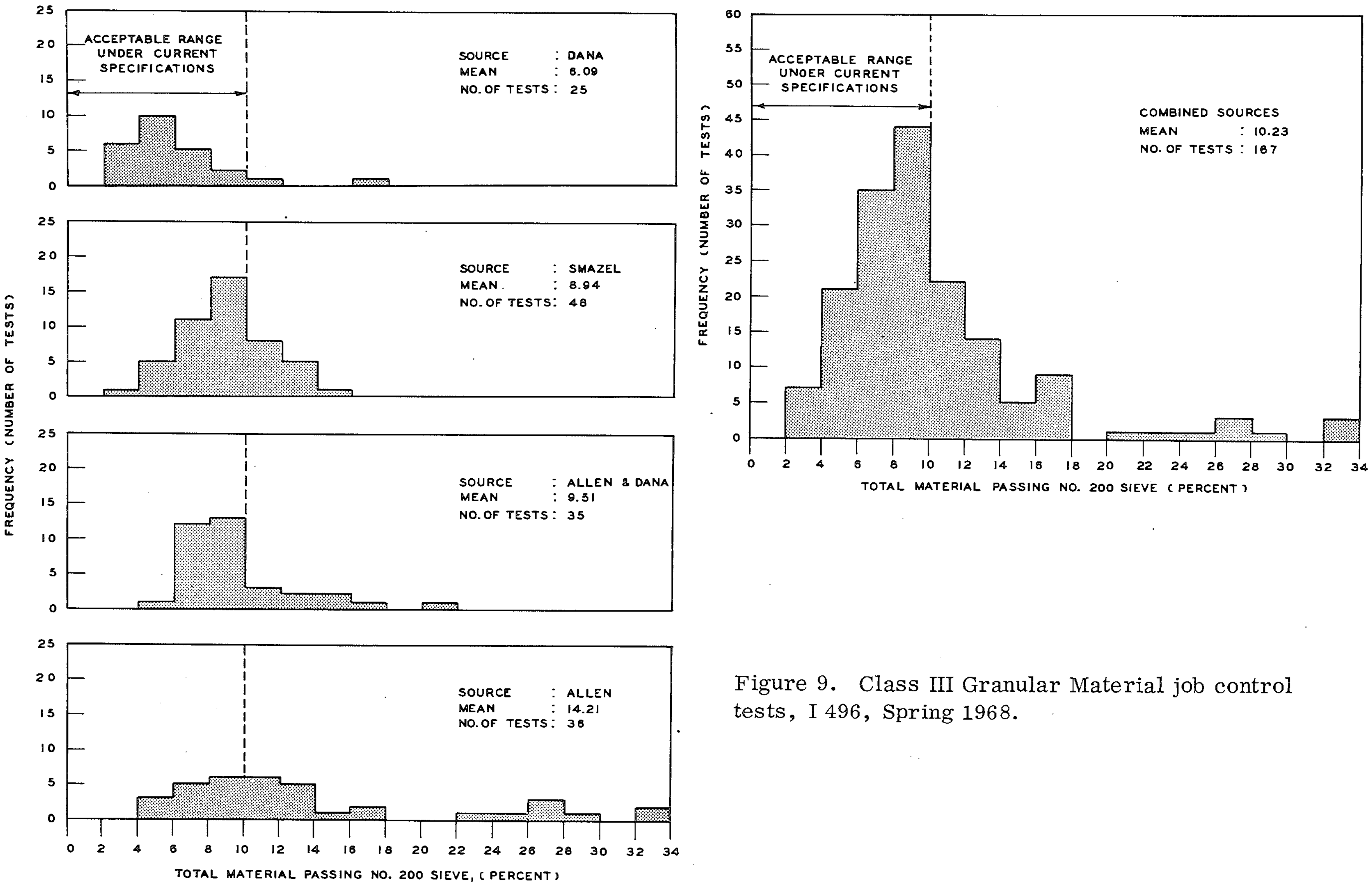


Figure 9. Class III Granular Material job control tests, I 496, Spring 1968.

It was decided that the plan should be designed to almost always reject material whose percent loss by washing was 15 percent or greater.

Figure 10 is an operating characteristic curve for the suggested sampling plan. This curve shows that the sampling plan reduces to only 10 percent the probability of accepting material with a loss by washing of 15 percent. As the loss by washing value increases, the chance of accepting further decreases. The plan can also be adjusted if these limits are not acceptable.

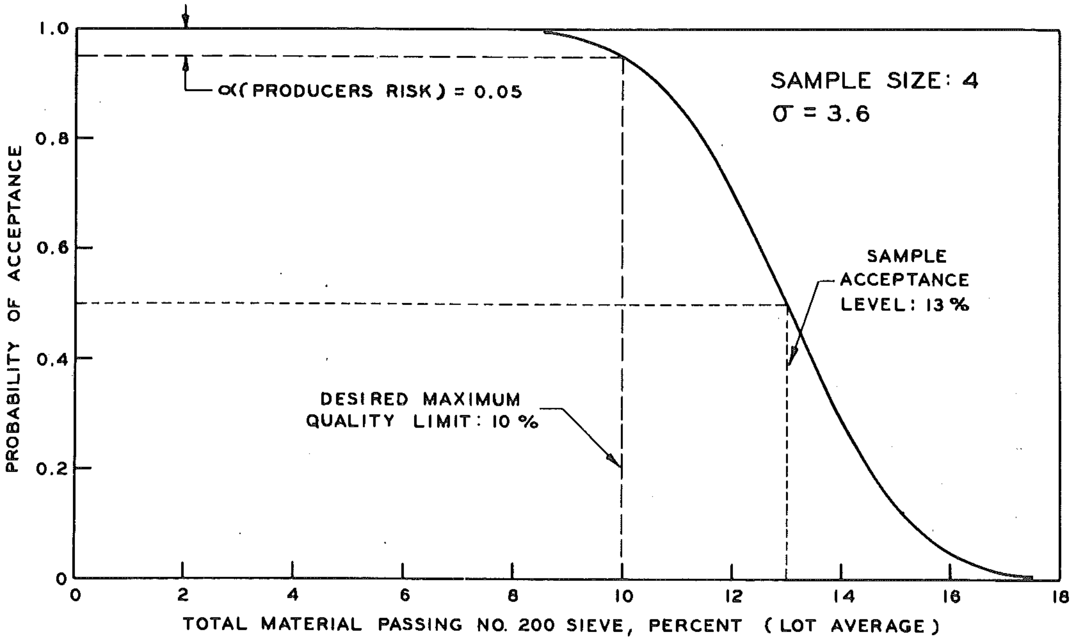


Figure 10. Operating characteristic curve of sampling plan for Class III Granular Material.

ACCEPTANCE TEST PLAN FOR LOSS BY WASHING FOR GRANULAR MATERIAL, CLASS III

a) Four routine test samples shall be taken from each lot. A lot to be considered one day's production from a given source or a maximum of 10,000 cu yd of material.

b) Each routine test sample shall be taken from an individual truckload selected in accordance with a random sampling plan.

c) Each test sample shall consist of about 25 lb and shall be composed of not less than three increments taken at random from the selected truckload of aggregate.

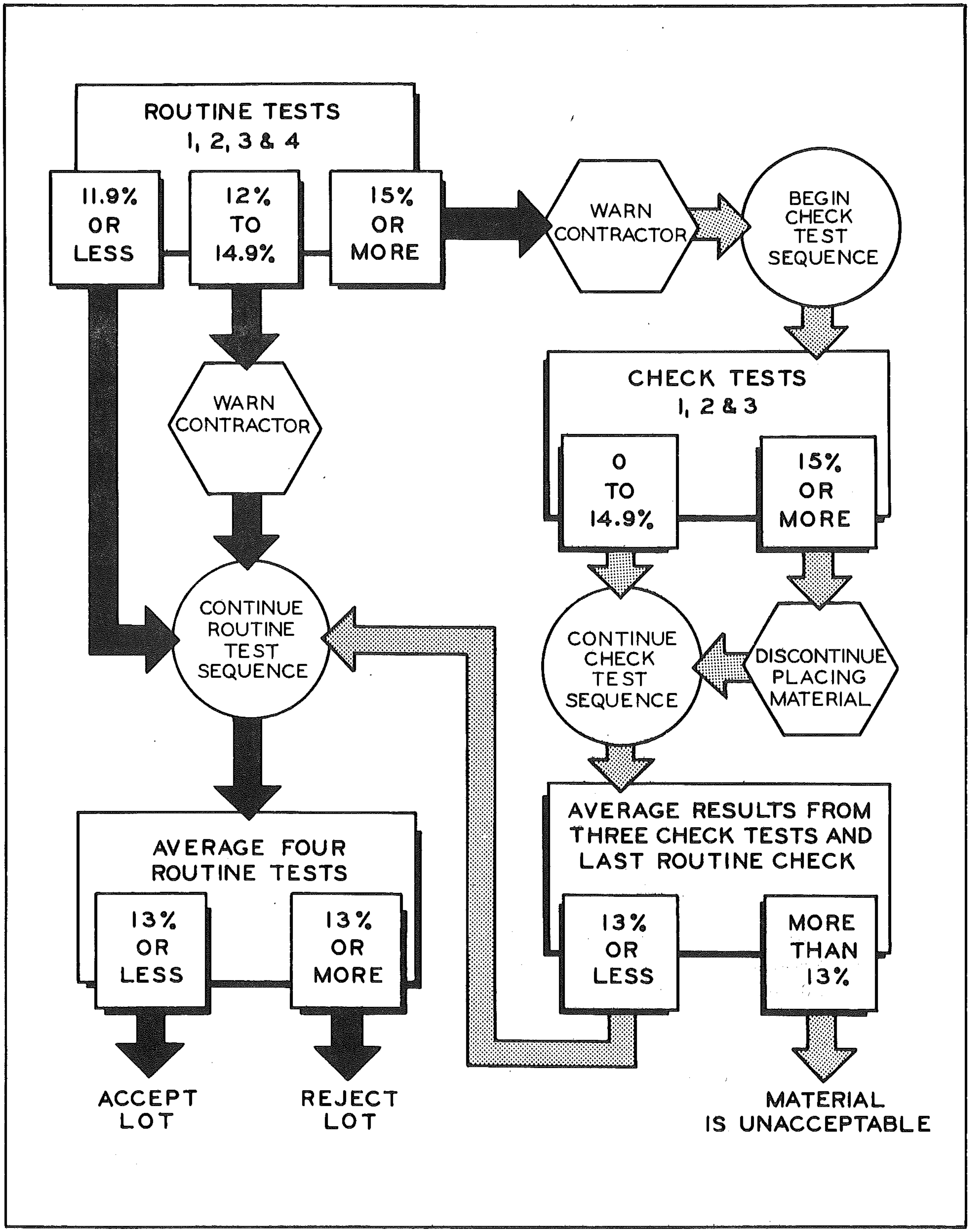


Figure 11. Test procedure for loss by washing Class III Granular Materials.

d) Individual truckloads will be sampled at the point of use. Material will be dumped from the truck to form a pile of uniform height of about 12 in.

e) The four routine test samples and all required check samples shall be individually tested in accordance with the provisions of AASHO Method T-27 using sieves with square openings. For loss by washing, material shall be tested in accordance with the provisions of AASHO Method T-11.

f) When check tests are required, test portions shall be taken as nearly as possible from successive trucks.

g) Testing procedures shall be in accordance with the provisions of the attached drawing (Fig. 11).

h) Limits for loss by washing shall be as follows:

Lot Tolerance - 15 percent
Desired Upper Avg. Quality Limit - 10 percent
Sample Acceptance Limit (Avg. of four tests) - 13 percent

In acceptance testing based on variables, that is, where a quality characteristic is actually measured, material variability makes it necessary to use three different limits rather than only one limit as used in older testing procedures.

These limits are defined as follows:

Lot Tolerance (LT) - Material with fines exceeding this proportion is considered to be definitely undesirable. Therefore, the sampling plan is designed so the probability of accepting material with a fines content exceeding the LT is minimal.

Desired Upper Average Quality Limit (DUL) - Material containing a proportion of fines less than or equal to this value is most desired. The sampling plan is designed to insure that almost all the material with an average fines content equal to or lower than the DUL will be accepted. Material with a fines content above the DUL will be rejected with a probability that rapidly increases between the DUL and the LT.

Sample Acceptance Limit (SAL) - Values from the specified number of samples from each lot are averaged. Material is rejected if the average value exceeds the SAL. This insures that the desired DUL and SAL are being achieved.

PROPOSED 22A AGGREGATE GRADATION FIELD EXPERIMENT

Under Phase I of this study, reported in Appendix A, a limited field experiment was conducted involving gradation inspection of 22A aggregate. The Phase I study provided much valuable information but was a short-term investigation limited to one aggregate source. Therefore, it was proposed that an expanded field study be made over a time span of at least 30 days and involving at least three different aggregate sources. Although the Research Laboratory believed that the study would provide a great deal of additional valuable information, the Advisory Committee vetoed the proposal as being a duplication of effort. The proposal, however, as originally presented to the Committee is outlined in this section.

Current 22A Aggregate Inspection Practices

Aggregate 22A for base course construction may consist of crushed stone, crushed gravel, or blast furnace slag, which conforms to the limits of grading, crushed material content, and abrasion resistance required by MDSH specifications. Specifications require a 25 percent minimum of crushed material for gravel and, based on AASHO T-4, 20 and 30 percent maximum wear for crushed and uncrushed gravel, respectively, and 30 percent maximum wear for stone. These limits are specified because of their effects on stability and abrasion resistance of the aggregate. Gradation or particle size distribution of an aggregate is determined by sieve analysis. Standard sieves (with square openings) for 22A aggregate required by current specification limits are as follows:

Sieve Size	Percent Passing
1 inch	100
3/4 inch	90-100
3/8 inch	65-86
No. 8	30-50
No. 200 (loss by washing)	3-7

Grading limits and maximum size are specified because of their effects on size of aggregate voids, degradation and permeability, frost action, segregation, and economy. Gradation specifications for 22A aggregate require that all tests from representative samples fall within the specified limits. A representative sample is one which, in the opinion of the inspector, represents an average condition of the material being sampled.

When acceptance of 22A aggregate is based upon visual and sampling inspection at the site of production or at the project, the trained aggregate inspector--under the supervision of the District Materials Supervisor--insures that the aggregate materials meet specifications and that proper methods of handling and stockpiling are used. He becomes familiar with plant processing and production problems and records the characteristics and location of the materials. If he gets a sample that does not meet specifications, he notifies the producer and tests another sample from the following production. If this sample still does not meet specifications, stockpiled material represented by the two faulty samples is rejected. If the result from either test falls within the specified limits, the material is accepted. MDSH specifications require one complete gradation analysis for each two hours of plant operation. Four or five tests per day will cover the production from an average gravel plant producing about 1,500 tons per 8-hr day. The inspector takes a representative sample by gathering material from different areas of the stockpile and combining it into a composite sample of about 60 to 80 lb. When the producer increases the production rate of 22A aggregate and the field inspector is unable to test the increased number of samples, he reports at once to his supervisor, who decides what action is to be taken.

Each composite or average sample is reduced by a Gilson Sample Splitter to a size (about 4,000g) suitable for testing for loss by washing (or passing the No. 200 sieve) and sieve analysis.

Proposed Field Study

Based on statistical concepts, a field study of gradation analysis has been designed to achieve the following objectives:

1) To establish the relative performance between the existing inspection practices and the suggested acceptance inspections based on random sampling.

2) To estimate variance components introduced into screening results by changes in aggregate materials, sampling, and testing procedures.

3) To further develop practical and meaningful acceptance limits for sieve analysis of 22A aggregate.

The proposed research project will be divided into three stages. The first stage is to develop a sampling plan to provide data that realistically represent the desired quality of the aggregate submitted for acceptance.

This will require proper recording and appraisal of the random data collected. This information, compared with that obtained by regular field inspectors, should disclose the relative difference between the two sampling procedures. The second stage is to evaluate the effects of changes in aggregate materials and in sampling and testing procedures. Well-trained inspectors assigned to this project must faithfully follow the instructions given in the proposed plan. The third stage is to compare the random sampling results of the first two stages with those obtained by current inspection practices. This should provide a firm ground for designing a practical and economical specification for sieve analysis of 22A aggregate. It should also provide some guidance for specifying gradation tolerances of other surfacing aggregates.

This study is not intended to: a) estimate process control of gravel plants, nor relative efficiency among gravel producers; b) evaluate relative performance among aggregate inspectors nor relative effects of different sample size, sampling, and testing equipment.

Research Procedure

All routine job control tests will be carried out in the normal manner. Additional tests recommended in this study will be carried out by special personnel without interference with normal job control. The suggested testing plan includes the following considerations:

A - Size and Location of Experimental Projects

In selecting the test site, it is important to consider different aggregate sources, contractor's procedures--including materials control, handling and stockpiling methods -- considered representative throughout the State. In addition, the testing program requires:

1) At least three similar projects with average gravel plants producing about 1,500 tons per 8-hr day to establish the desired statistical parameters.

2) At least 100 random duplicated samples per each project to obtain reliable results.

3) Studies on projects or jobs with inexperienced contractors or with unusual materials should be avoided.

B - Controlled Variables

Those factors that must be known and recorded during the experiment include the aggregate source, type of commercial plant, production methods, and control procedures; including type of equipment for handling and stockpiling the finished product.

C - Type of Tests

Random samples from flat-layered stockpiles will be used to determine gradation and crushed material of 22A aggregate. Standard sieves to be used are 1 in., 3/4 in., 3/8 in., No. 8, and No. 200. Gradation tests are to be determined by the current method, AASHO T-27. Loss by washing is to be determined by AASHO T-11, currently applied to aggregate material finer than the No. 200 sieve. Grading results are to be reported to one decimal place on standard form Nos. 1900 and 1901. Sieves are to be calibrated before being used by research personnel. Furthermore, the No. 8 and No. 200 sieves are to be periodically calibrated in accordance with E 11. Sieves compared with the Standard Sieve, should agree within 5 percent. Sieves with deviations exceeding 10 percent should not be used in this experiment.

Percentage of crushed material is to be determined as specified for surfacing aggregate, Article 7.02.02 of the Michigan 1967 Standard Specifications.

D - Sampling Procedure

The significance of the proposed experiment depends on how objectively and consistently the random sampling plan is applied during continuous production of 22A aggregate. Sampling and testing are to be conducted by well-trained inspectors under the supervision of the Research Laboratory. Sampling and testing for the study will be in addition to, and not interfere with, job control carried on in the conventional manner by regular aggregate inspectors.

A random sample (or probability sample) is one in which each increment of material from a lot has an equal chance of being included in the sample. A lot is defined as "...a day's production of the same aggregate material from the same source, produced under the same operating conditions and stockpiled according to a specified construction procedure."

Regardless of the sampling procedure being used, the aggregate inspector knows that successive samples taken from the same stockpile are

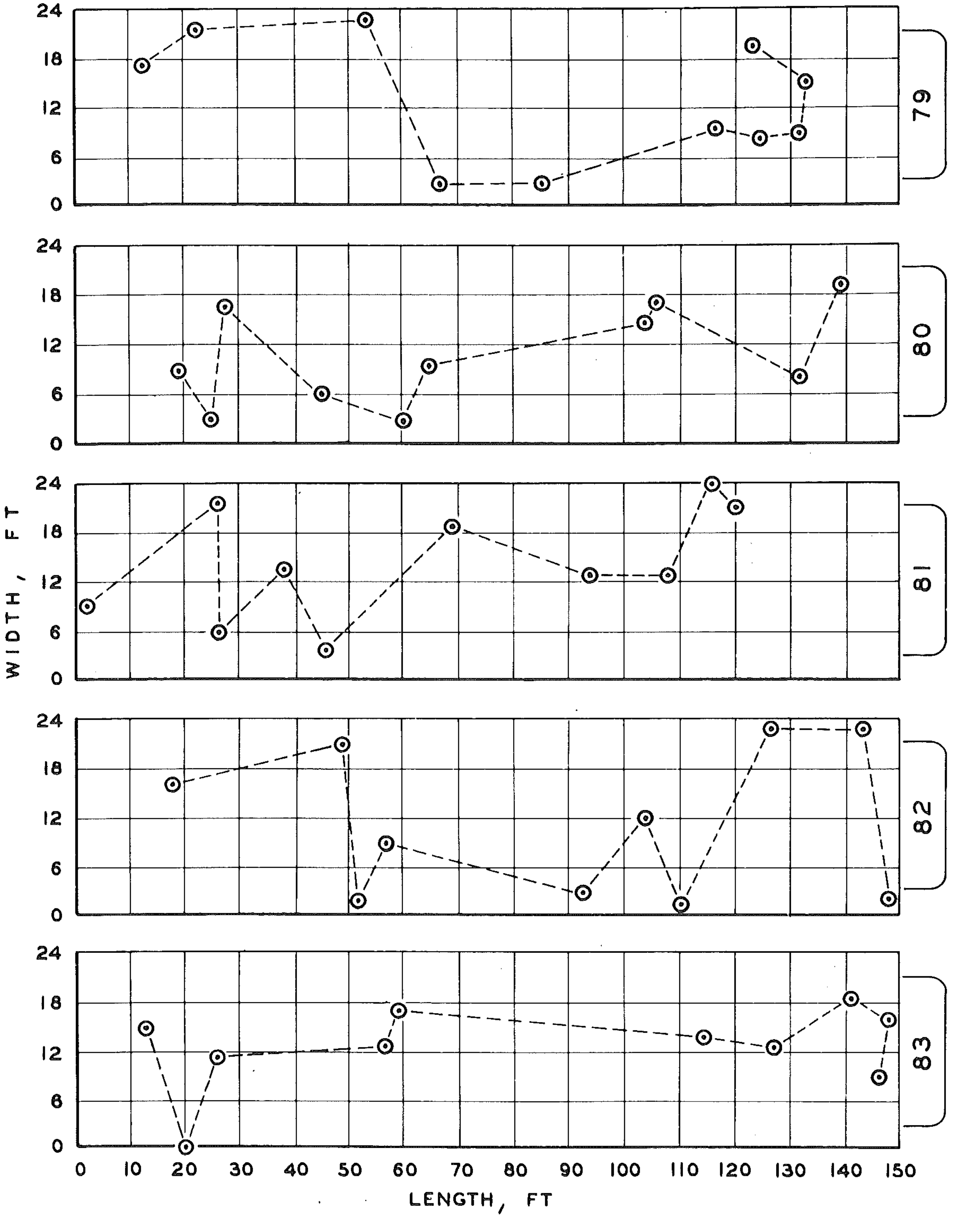


Figure 12. Plan view of random sampling layouts for flat-layered stockpiles (approximately 150 ft by 24 ft) with 10 sampling locations for selecting a composite sample of 22A aggregate.

usually different. He also knows that aggregate materials tend to segregate and that, despite sampling variations and segregation, his problem is to make correct inferences about the quality of the aggregate source from which the sample is drawn. Thus, it is extremely important when making up a random sample, that each increment of stockpiled material have an equal likelihood of being included in the composite sample. In this connection, it is difficult to apply the probability concept whenever increments of materials cannot be taken from the interior of the stockpile. However, the probability concept may be approximated by selecting, at random, increments of materials from flat horizontal layers with thickness not exceeding 6 in. in depth.

Increments of random sampling locations shall be selected from 100 different sampling layout cards designed by the Research Laboratory. Five typical layout cards are shown in Figure 12. When a flat-layered stockpile is to be sampled (approximately 150 ft by 24 ft), the inspector will draw at random a three-digit number from Table 5. Suppose this number is 081; the last two digits (81) correspond to the number of the layout to be used. Layout No. 81 will be used to determine ten locations for sampling. Each inspector will use two bags for gathering duplicate samples. At each location, one scoop (about 6 lb) of gravel will be placed into each bag. Thus, each bag will contain about 60 lb of gravel as a complete sample. Each composite sample will be reduced by a Gilson Sample Splitter to approximately 4,000g to determine loss by washing (or passing No. 200 sieve) and for a sieve analysis (Fig. 13).

Figure 14 shows the sequence of testing operations. A minimum of two duplicate samples (a total of four samples) are taken daily, at random, from 6-in. deep flat layers (Fig. 12) being stockpiled from a continuous production of 22A aggregate. Layers to be sampled will be selected at random using Table 5. It is estimated that each layer will weigh about 117 tons. Two samples are sent to the laboratory to be tested. The other two samples are immediately tested in the field, using current equipment and testing procedures (Fig. 13). The random sampling plan is continued until the aggregate production is ended, or until a minimum of 200 gradation tests (representing 100 samples) are run in the field from each aggregate source under study.

Minimum requirements for conducting the investigation are listed below. If all requirements cannot be fulfilled, the investigation will be modified as necessary. Requirements are:

1) At least three different gravel pits producing about 1,500 tons per 8-hr day.

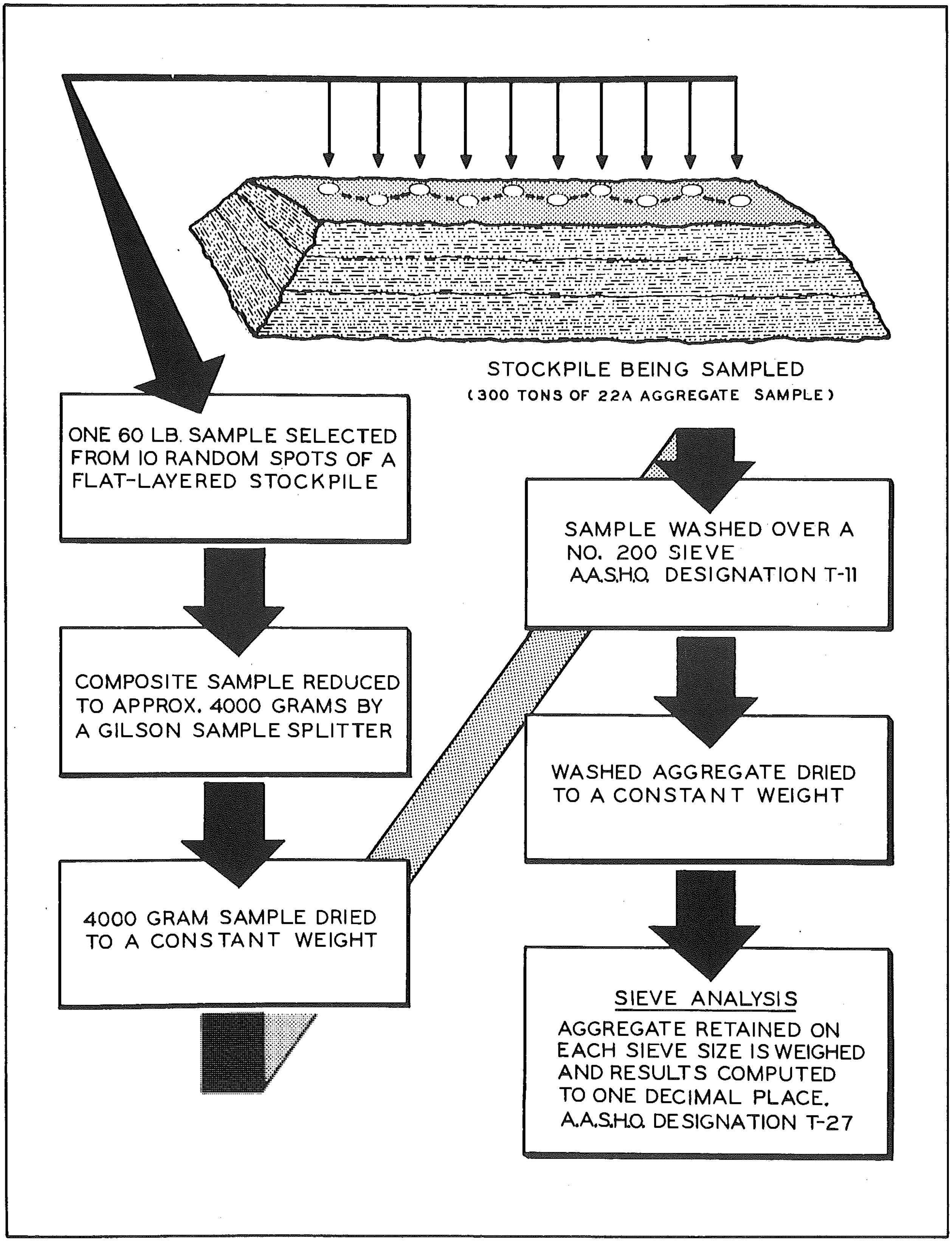


Figure 13. Flow chart for sampling and screening coarse aggregates.

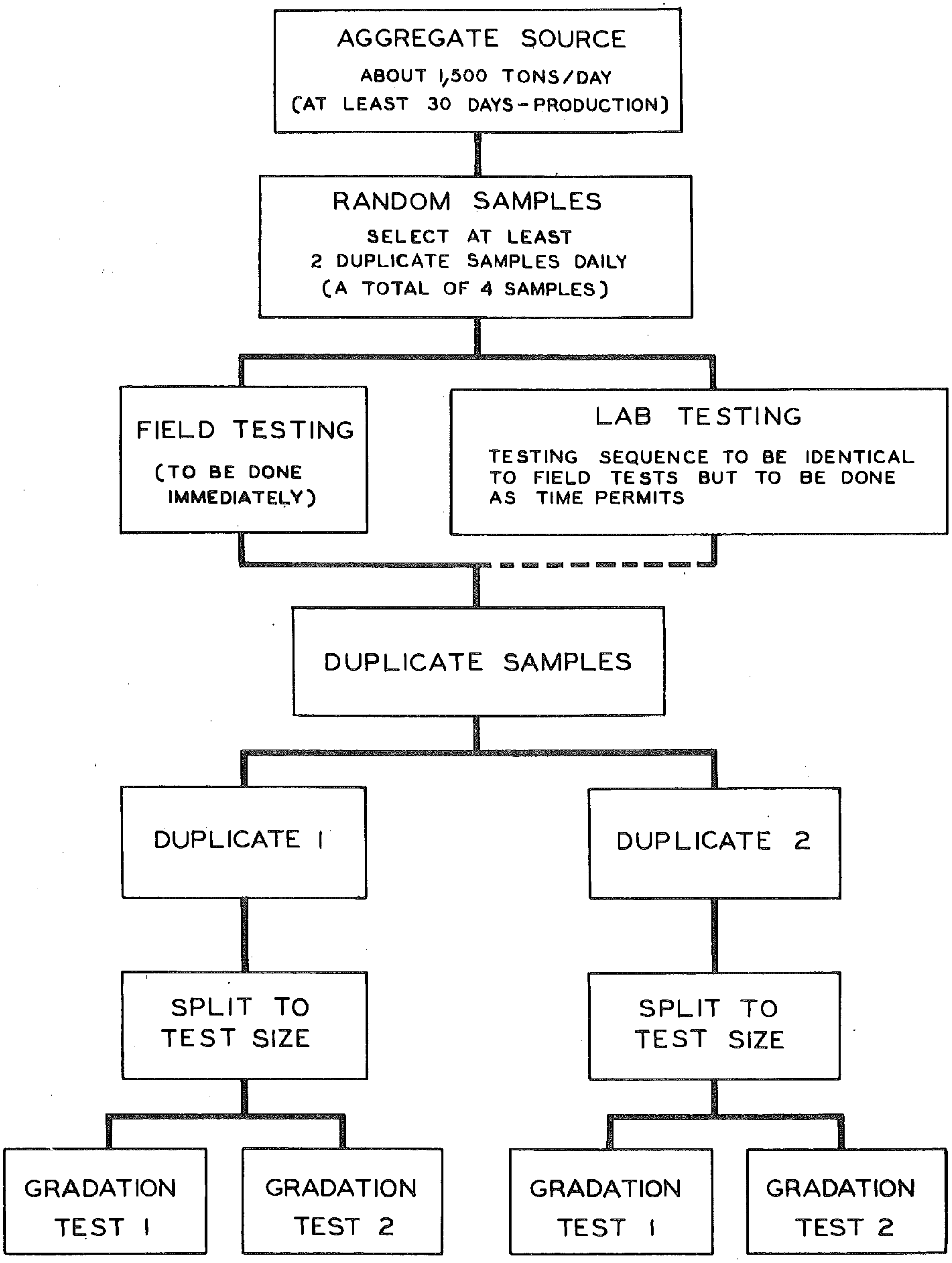


Figure 14. Controlled experiment for gradation analysis of 22A aggregate.

TABLE 5
RANDOM NUMBERS

.576	.730	.430	.754	.271	.870	.732	.721	.998	.239
.892	.948	.858	.025	.935	.114	.153	.508	.749	.291
.669	.726	.501	.402	.231	.505	.009	.420	.517	.858
.609	.482	.809	.140	.396	.025	.937	.310	.253	.761
.971	.824	.902	.470	.997	.392	.892	.957	.640	.463
.053	.899	.554	.627	.427	.760	.470	.040	.904	.993
.810	.159	.225	.163	.549	.405	.285	.542	.231	.919
.081	.277	.035	.039	.860	.507	.081	.538	.986	.501
.982	.468	.334	.921	.690	.806	.879	.414	.106	.031
.095	.801	.576	.417	.251	.884	.522	.235	.398	.222
.509	.025	.794	.850	.917	.887	.751	.608	.698	.683
.371	.059	.164	.838	.289	.169	.569	.977	.796	.996
.165	.996	.356	.375	.654	.979	.815	.592	.348	.743
.477	.535	.137	.155	.767	.187	.579	.787	.358	.595
.788	.101	.434	.638	.021	.894	.324	.871	.698	.539
.566	.815	.622	.548	.947	.169	.817	.472	.864	.466
.901	.342	.873	.964	.942	.985	.123	.086	.335	.212
.470	.682	.412	.064	.150	.962	.925	.355	.909	.019
.068	.242	.667	.356	.195	.313	.396	.460	.740	.247
.874	.420	.127	.284	.448	.215	.833	.652	.601	.326
.897	.877	.209	.862	.428	.117	.100	.259	.425	.284
.875	.969	.109	.843	.759	.239	.890	.317	.428	.802
.190	.696	.757	.283	.666	.491	.523	.665	.919	.146
.341	.688	.587	.908	.865	.333	.928	.404	.892	.696
.846	.355	.831	.218	.945	.364	.673	.305	.195	.887
.882	.227	.552	.077	.454	.731	.716	.265	.058	.075
.464	.658	.629	.269	.069	.998	.917	.217	.220	.659
.123	.791	.503	.447	.659	.463	.994	.307	.631	.422
.116	.120	.721	.137	.263	.176	.798	.879	.432	.391
.836	.206	.914	.574	.870	.390	.104	.755	.082	.939
.636	.195	.614	.486	.629	.663	.619	.007	.296	.456
.630	.673	.665	.666	.399	.592	.441	.649	.270	.612
.804	.112	.331	.606	.551	.928	.830	.841	.602	.183
.360	.193	.181	.399	.564	.772	.890	.062	.919	.875
.183	.651	.157	.150	.800	.875	.205	.446	.648	.685

2) At least 30 days of aggregate production for each gravel pit.

3) At least three laboratory aides to perform the required field tests.

4) For each project, regular testing equipment as follows:

1	Gilson Sample Splitter
1	set coarse aggregate sieves
1	set fine aggregate sieves
4	washing pans
5	burners
1	100-lb gas tank per week
1	trowel, counterbrush, scoop, spoon, round file, sample pail, screenbrush, slide rule, mechanical analysis book, daily aggregate reports, envelopes, stamps, pencils
1	balance with pans, set of weights, 1 g to 1 kg, extra kilogram weights.
60	sample sacks.

V
RECOMMENDATIONS AND CONCLUSIONS

Organization for Quality Control

This research study has been active for several years and has resulted in the development of a number of statistical acceptance testing plans; it has not produced any significant changes in Michigan's quality control program. Among the apparent reasons for the refusal to adopt any significant statistically based plans are the emphasis on statistical methods which are unfamiliar to most Department personnel, plus a general resistance to leaving the security of familiar methods for something strange and not fully comprehended. However, these objections to the program might be overcome if sufficient and competent effort were put into educating key people and selling the concept of statistical control.

There is currently no formal quality control organization in the Department. Rather, scattered personnel--having other competing responsibilities--develop proposals and plans but haven't the time nor opportunity to promote them.

Recommendations

Large industries reach a point in their growth when they must formally organize a quality control program. Each industry, because of its own unique problems, must tailor their quality control organization to suit their individual needs. Such quality control organizations function only to develop and implement quality control methods and, to be effective, they must have the support of higher management.

If this Highway Department is to implement an effective quality control program, it too must develop and support a separate quality control group. The nucleus for such a group could involve an engineer and a statistician, knowledgeable and experienced in statistical quality control methods, and could be supplemented with persons having considerable field experience. This group could be developed within an existing Departmental Section or Division, but its support would have to come directly and distinctly from higher management.

Such a quality control group would serve in a staff capacity and would have authority to define, coordinate, and standardize Departmental job control and acceptance testing procedures and methods. It would not assume responsibility for all quality control or assurance functions now carried on in various divisions. However, the group would have the obligation and authority to review and make recommendations for all programs dealing with quality of materials, finished products, or controlled experiments dealing with quality.

Specific duties of the group would include:

1) Design sampling plans for general use and specific situations.
2) Design means for measuring accuracy and quality of inspection work.
3) Design and recommend experiments to secure maximum information about a quality characteristic.
4) Draft Departmental policy on quality.
5) Conduct quality audits and surveys.
6) Prepare periodic reports on quality performance.
7) Design and conduct appropriate training programs, as required, for all levels of personnel.
8) Provide consultation on statistical methods for quality control.
9) Review specifications and test procedures.

Conclusions

Although no portion of this study has been incorporated into the Department's operating procedures, experience has been gained which should be valuable as statistical methods are adopted in the future. Objective 1 was met to a partial extent, for it was felt that some areas were found to be amenable to the application of quality control techniques; these are listed in the Introduction, and three areas are specifically treated in Sections II, III, and IV. In these three areas, Objective 2 was met to the extent that acceptance plans were developed. Objective 2, the development of suitable acceptance plans, could not be pursued to the extent that the plans could be refined since none of them were actually incorporated into the Departmental procedure. Due to this fact, Objective 3, rewriting existing specifications in the light of these plans, could not be fulfilled--although tentative suggestions are given in the three areas specifically treated as to possible revisions. The final Objective, number 4, could not be attempted since the inertia evidenced toward the proposed plans precluded the writing of a Departmental Manual.

It was felt, however, that treatment of the three specific areas, modulus-of-rupture, neoprene joint seal, and 22A aggregate, did show the value and practicability of statistical testing, despite the fact that management could not be convinced sufficiently to change existing procedures, or even allow the expansion of pilot programs in these areas.

A steering committee consisting of persons unfamiliar with statistical quality control has proven unwieldy. As suggested in the Recommendations, it is felt that if expertise is needed in a particular area, it should be obtained as needed by the quality control group.

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the Federal Highway Administration.

APPENDIX

HIGHWAY QUALITY CONTROL PROGRAM
Statistical Parameters

Progress Report on a Highway Planning and Research Project
Conducted in Cooperation With
The U. S. Department of Commerce--Bureau of Public Roads

Research Laboratory Section
Office of Testing and Research
Research Project 63 G-123
Research Report No. R-572

State of Michigan
Department of State Highways

Lansing, March 1966

INTRODUCTION

This report has been prepared to summarize statistical analyses that have been conducted on various highway materials and construction procedures. This summary is in accordance with a Bureau of Public Roads memorandum from F. B. Farrell, Bureau of Public Roads Regional Engineer, to Division Engineers in Indiana, Michigan, and Wisconsin on January 13, 1966, titled "Statistical Parameters - Quality Control Program." The information gathered to date on Phases I and II of the Department's HPR program on Highway Quality Control has been screened to present all available data requested in this memorandum.

The data in this report arise from two sources: 1) random sampling of past field construction records of job control testing, and 2) a statistically designed experiment on field testing of aggregate gradation conducted to determine, if possible, specific causes for testing variation. For the first set of data, it was possible to obtain only the following statistical parameters requested in the Bureau's memorandum: the number of test measurements, the arithmetic mean and overall variance, and the overall standard deviation. Thus, specific causes for variation cannot be assigned. For the second set of data, deriving from the experimental program, some information was available on sampling and testing variance due to different inspectors and different equipment (screening sieves). However, the experimental design was not set up to differentiate between material and experimental variance. In every case, the specification requirements are illustrated or discussed along with the data, and therefore, the separate transmittal of specifications suggested in the Bureau's memorandum is deemed unnecessary.

AGGREGATE GRADATION— RECORDS SURVEY

The chief sources of information were daily job control records of aggregate field testing covering the period 1959 through 1961. About 120 highway projects were classified by finding, counting, and identifying daily records; types of aggregate; and producers. The types of aggregate materials involved in this study were limestone, gravel, and slag. Detailed analysis of the data established proper sampling procedures and the computations required for appropriate statistical evaluation.

It should be noted that the data analyzed and shown in Table 1 did not include the percents passing for all sieve size requirements of the specifications, but did include representative sieve sizes for each aggregate. In addition, it should be noted that Table 1 consists exclusively of data from accepted batches of material. The 4A, 6A, 6AA, and 10A aggregates are used for coarse aggregate in portland cement concrete for either pavements or structures; the 9A aggregate for bituminous base and binder courses; 20A for bituminous aggregate surface courses; 22A for aggregate base and surface courses, aggregate shoulders, and aggregate resurfacing; 23A for aggregate shoulders and aggregate approaches; 24A for selected subbase; 25A for bituminous concrete wearing course; 26B for bituminous non-skid surface treatment; and 2NS sand for portland cement concrete. These materials were produced in accordance with then current Michigan specifications; however, in the 1965 edition of the Michigan Standard Specifications, certain revisions were made which changed the requirements for some of these aggregates, and two classifications (10A and 26B) no longer exist.

Since the Table 1 data consisted only of results from accepted batches, a study was subsequently made to determine the proportion of accepted-to-rejected material, and how the inclusion of data from the rejected material would influence the overall distribution. Aggregate type 22A was selected for this study since from experience it was noted that this type generally has a larger proportionate amount of rejected material than other aggregate types. Data from 11 projects were studied with the results summarized in Table 2. Approximately 3.7 percent of the material tested was rejected for one or more deviations from specification requirements. For the 104 rejected results, 42 (or 40.4 percent) failed to pass the No. 10 sieve within requirements, 22 (or 21.1 percent) failed to meet requirements on loss by washing (No. 200 sieve), 8 (or 7.7 percent) failed to meet 3/8-in. sieve requirements, while 16 (or 15.4 percent) failed to meet the requirements for more than one sieve, and 16 (or 15.4 percent) failed for other reasons than gradation.

TABLE 1
SUMMARY OF GRADATION TESTS ON ACCEPTED MATERIAL OF DIFFERENT AGGREGATE TYPES

Material and Producer	Item	Type 4A						
		Sieve Size and Spec. Limits, %						Loss by Washing 0.8 % max.
		2-1/2 in. 100	2 in. 95-100	1-1/2 in. 65-90	1 in. 10-40	1/2 in. 0-20	3/8 in. 0-5	
Gravel Producer A (300 Tests)	Avg. Gradation,%	Not Analyzed	Not Analyzed	77.8	22.2	2.2	1.1	Not Analyzed
	Std. Deviation,%			8.0	8.1	1.8	1.0	
	% of Tests:							
	Within Specs., %			95.3	97.0	100.0	99.3	
	Above,%			0.7	2.0	0.0	0.7	
	Below,%			4.0	1.0	0.0	0.0	
Gravel Producer B (300 Tests)	Avg. Gradation,%			80.6	26.1	3.7	1.4	
	Std. Deviation,%			7.0	7.0	3.7	1.5	
	% of Tests:							
	Within Specs., %			97.6	95.7	100.0	96.7	
	Above,%			1.7	3.3	0.0	3.3	
	Below,%			0.7	1.0	0.0	0.0	
Gravel Producer C (300 Tests)	Avg. Gradation,%			74.7	23.8	0.7		
	Std. Deviation,%			6.4	8.1	1.1		
	% of Tests:							
	Within Specs., %			96.0	98.0	100.0	100.0	
	Above,%			1.0	1.7	0.0	0.0	
	Below,%			3.0	0.3	0.0	0.0	
Gravel Producer D (300 Tests)	Avg. Gradation,%			75.5	29.8	5.2	1.9	
	Std. Deviation,%			7.9	8.8	3.8	1.7	
	% of Tests:							
	Within Specs., %			92.3	93.0	99.7	94.3	
	Above,%			1.7	7.0	0.3	5.7	
	Below,%			6.0	0.0	0.0	0.0	
Gravel Producer E (300 Tests)	Avg. Gradation,%			80.2	22.9	5.1	2.4	
	Std. Deviation,%			7.1	7.0	4.2	2.1	
	% of Tests:							
	Within Specs., %			95.0	97.4	99.0	92.3	
	Above,%			3.3	2.3	1.0	7.7	
	Below,%			1.7	0.3	0.0	0.0	

Material and Producer	Item	Type 4A (Cont.)						
		Sieve Size and Spec. Limits, %						Loss by Washing 1.5 % max.
		2-1/2 in. 100	2 in. 95-100	1-1/2 in. 65-90	1 in. 10-40	1/2 in. 0-20	3/8 in. 0-5	
Stone Producer F (300 Tests)	Avg. Gradation,%	Not Analyzed	Not Analyzed	74.0	28.7	1.0	0.7	Not Analyzed
	Std. Deviation,%			5.9	6.2	0.9	0.6	
	% of Tests:							
	Within Specs., %			99.7	99.7	100.0	100.0	
	Above,%			0.0	0.3	0.0	0.0	
	Below,%			0.3	0.0	0.0	0.0	
Stone Producer G (300 Tests)	Avg. Gradation,%			77.2	25.9	3.2	1.7	
	Std. Deviation,%			7.6	8.6	4.6	1.7	
	% of Tests:							
	Within Specs., %			97.0	96.7	100.0	95.0	
	Above,%			0.7	2.3	0.0	5.0	
	Below,%			2.3	1.0	0.0	0.0	
Stone Producer H (300 Tests)	Avg. Gradation,%			77.4	19.6	2.9	1.7	
	Std. Deviation,%			7.2	7.3	2.6	2.5	
	% of Tests:							
	Within Specs., %			94.7	95.4	100.0	94.0	
	Above,%			2.0	0.3	0.0	6.0	
	Below,%			3.3	4.3	0.0	0.0	
Slag Producer I (300 Tests)	Avg. Gradation,%			76.1	21.8	1.9	1.3	
	Std. Deviation,%			7.0	4.7	1.2	0.8	
	% of Tests:							
	Within Specs., %			96.3	99.7	100.0	100.0	
	Above,%			0.0	0.0	0.0	0.0	
	Below,%			3.7	0.3	0.0	0.0	

NOTE: Method of sampling was by proportional allocation of daily job control records of materials which were accepted. Sizes of samples as indicated in first column for each material and producer. Current Specifications (1965 Edition) are different than those shown, as follows:

4A--Unchanged
6A--Changed
6AA--Changed
9A--Changed
10A--No Longer Exists
20A--Changed
22A--Changed Somewhat
23A--Changed Slightly
24A--Changed
25A--Changed Significantly
26B--No Longer Exists
2NS--Unchanged

TABLE 1 (cont.)

SUMMARY OF GRADATION TESTS ON ACCEPTED MATERIAL OF DIFFERENT AGGREGATE TYPES

Material and Producer	Item	Type 6A						
		Sieve Size and Spec. Limits, %						Loss by Washing 0.8 % max.
		2-1/2 in. 100	2 in. 95-100	1-1/2 in. 90-100	1 in. 60-90	1/2 in. 25-55	No. 4 0-8	
Gravel Producer A (300 Tests)	Avg. Gradation,%			99.8	77.3	37.3	0.9	
	Std. Deviation,%			0.7	6.3	6.9	0.7	
	% of Tests:							
	Within Specs., %			100.0	99.3	99.3	100.0	
	Above,%			0.0	0.7	0.0	0.0	
	Below,%			0.0	0.0	0.7	0.0	
Gravel Producer U (226 Tests)	Avg. Gradation,%	Not Analyzed	Not Analyzed	99.9	79.1	34.5	0.8	Not Analyzed
	Std. Deviation,%			0.4	6.4	7.1	0.8	
	% of Tests:							
	Within Specs., %			100.0	99.1	95.6	100.0	
	Above,%			0.0	0.9	0.9	0.0	
	Below,%			0.0	0.0	3.5	0.0	
Gravel Producer V (222 Tests)	Avg. Gradation,%			98.6	72.4	33.7	2.5	
	Std. Deviation,%			1.6	7.5	7.4	1.6	
	% of Tests:							
	Within Specs., %			100.0	96.8	94.6	100.0	
	Above,%			0.0	0.9	0.0	0.0	
	Below,%			0.0	2.3	5.4	0.0	

Material and Producer	Item	Type 6AA				
		Sieve Size and Spec. Limits, %				Loss by Washing 0.8 % max.
		1-1/2 in. 100	1 in. 95-100	1/2 in. 30-60	No. 4 0-8	
Gravel Producer S (89 Tests)	Avg. Gradation,%	Not Analyzed	99.9	47.4	3.5	Not Analyzed
	Std. Deviation,%		0.5	6.3	1.6	
	% of Tests:					
	Within Specs., %		100.0	98.9	100.0	
	Above,%		0.0	1.1	0.0	
	Below,%		0.0	0.0	0.0	

Material and Producer	Item	Type 9A				
		Sieve Size and Spec. Limits, %				Loss by Washing 3 % max.
		1-1/4 in. 100	3/4 in. 45-65	3/8 in. 0-25	No. 4 0-10	
Crushed Gravel Producer A (147 Tests)	Avg. Gradation,%		54.8	4.5	1.0	
	Std. Deviation,%		5.5	3.1	0.6	
	% of Tests:					
	Within Specs., %		97.9	100.0	100.0	
	Above,%		0.7	0.0	0.0	
	Below,%		1.4	0.0	0.0	
Stone Producer H (108 Tests)	Avg. Gradation,%	Not Analyzed	51.6	8.3	3.0	Not Analyzed
	Std. Deviation,%		6.3	4.7	1.8	
	% of Tests:					
	Within Specs., %		90.7	100.0	100.0	
	Above,%		1.9	0.0	0.0	
	Below,%		7.4	0.0	0.0	
Crushed Gravel Producer T (80 Tests)	Avg. Gradation,%		52.7	3.0	1.6	
	Std. Deviation,%		5.5	1.9	1.0	
	% of Tests:					
	Within Specs., %		95.0	100.0	100.0	
	Above,%		0.0	0.0	0.0	
	Below,%		5.0	0.0	0.0	

TABLE 1 (cont.)
SUMMARY OF GRADATION TESTS ON ACCEPTED MATERIAL OF DIFFERENT AGGREGATE TYPES

Material and Producer	Item	Type 10A				
		Sieve Size and Spec. Limits, %				Loss by Washing 0.8 % max.
		1-1/2 in. 100	1 in. 95-100	1/2 in. 35-65	No. 4 0-8	
Gravel Producer A (300 Tests)	Avg. Gradation,%	Not Analyzed	98.6	47.8	0.9	Not Analyzed
	Std. Deviation,%		1.6	8.3	0.8	
	% of Tests:					
	Within Specs., %		98.7	96.3	100.0	
	Above,%		0.0	1.0	0.0	
	Below,%		1.3	2.7	0.0	
Gravel Producer B (300 Tests)	Avg. Gradation,%		99.2	50.5	2.2	
	Std. Deviation,%		1.3	7.5	2.2	
	% of Tests:					
	Within Specs., %		99.7	98.3	98.3	
	Above,%		0.0	0.7	1.7	
	Below,%		0.3	1.0	0.0	
Gravel Producer C (300 Tests)	Avg. Gradation,%		99.6	50.3	3.2	
	Std. Deviation,%		1.0	8.1	2.4	
	% of Tests:					
	Within Specs., %		99.7	96.3	98.3	
	Above,%		0.0	2.0	1.7	
	Below, %		0.3	1.7	0.0	
Gravel Producer D (300 Tests)	Avg. Gradation,%		97.9	46.4	1.8	
	Std. Deviation,%		1.7	7.4	1.8	
	% of Tests:					
	Within Specs., %		96.3	96.7	100.0	
	Above,%		0.0	0.3	0.0	
	Below,%		3.7	3.0	0.0	

Material and Producer	Item	Type 10A (Cont.)				
		Sieve Size and Spec. Limits, %				Loss by Washing 1.5 % max.
		1-1/2 in. 100	1 in. 95-100	1/2 in. 35-65	No. 4 0-8	
Gravel Producer E (300 Tests)	Avg. Gradation,%	Not Analyzed	99.0	47.2		Not Analyzed
	Std. Deviation,%		1.7	8.4		
	% of Tests:					
	Within Specs., %		97.0	93.4	100.0	
	Above,%		0.0	1.3	0.0	
	Below,%		3.0	5.3	0.0	
Stone Producer F (219 Tests)	Avg. Gradation,%		99.99	44.4	2.2	
	Std. Deviation,%		0.1	7.5	1.7	
	% of Tests:					
	Within Specs., %		100.0	100.0	98.6	
	Above,%		0.0	0.0	1.4	
	Below,%		0.0	0.0	0.0	
Stone Producer G (300 Tests)	Avg. Gradation,%		99.5	47.5	3.4	
	Std. Deviation,%		1.1	8.3	1.9	
	% of Tests:					
	Within Specs., %		99.7	95.7	98.0	
	Above,%		0.0	0.0	2.0	
	Below,%		0.3	4.3	0.0	
Slag Producer I (271 Tests)	Avg. Gradation,%		98.7	45.5		
	Std. Deviation,%		1.2	6.7		
	% of Tests:					
	Within Specs., %		99.6	98.2	99.6	
	Above,%		0.0	0.0	0.4	
	Below, %		0.4	1.8	0.0	

TABLE 1 (cont.)
SUMMARY OF GRADATION TESTS ON ACCEPTED MATERIAL OF DIFFERENT AGGREGATE TYPES

Material and Producer	Item	Type 20A					
		Sieve Size and Spec. Limits, %					Loss by Washing 0-5 %
		3/4 in. 100	3/8 in. 60-80	No. 10 40-50	No. 40 15-30	No. 200 0-5	
Gravel Producer K (140 Tests)	Avg. Gradation,%		75.0	44.2	22.6	3.1	
	Std. Deviation,%		4.7	3.0	3.7	0.7	
	% of Tests:						
	Within Specs., %		90.0	73.6	98.6	99.3	
	Above,%		10.0	10.7	1.4	0.7	
	Below,%		0.0	15.7	0.0	0.0	
Gravel Producer N (172 Tests)	Avg. Gradation,%	Not Analyzed	77.9	43.7	17.1	4.1	Not Analyzed
	Std. Deviation,%		2.5	3.3	2.6	0.6	
	% of Tests:						
	Within Specs., %		87.8	90.7	93.0	95.3	
	Above,%		12.2	8.1	0.0	4.7	
	Below,%		0.0	1.2	7.0	0.0	
Gravel Producer O (193 Tests)	Avg. Gradation,%		77.0	43.5	17.8	4.1	
	Std. Deviation,%		3.0	3.8	2.6	0.6	
	% of Tests:						
	Within Specs., %		91.7	85.5	96.4	96.4	
	Above,%		8.3	3.1	0.0	3.6	
	Below,%		0.0	11.4	3.6	0.0	

Material and Producer	Item	Type 22A				
		Sieve Size and Spec. Limits, %				Loss by Washing 3-7 %
		1 in. 100	3/4 in. 90-100	3/8 in. 65-85	No. 10 30-45	
Gravel Producer C (300 Tests)	Avg. Gradation,%		96.5	71.7	40.8	
	Std. Deviation,%		2.1	4.5	4.3	
	% of Tests:					
	Within Specs., %		100.0	94.7	87.3	
	Above,%		0.0	0.0	11.7	
	Below,%		0.0	5.3	1.0	
Gravel Producer J (300 Tests)	Avg. Gradation,%		95.5	72.8	39.6	
	Std. Deviation,%		2.6	4.6	3.7	
	% of Tests:					
	Within Specs., %		98.0	97.0	92.4	
	Above,%		0.0	0.0	7.3	
	Below,%		2.0	3.0	0.3	
Gravel Producer M (300 Tests)	Avg. Gradation,%	Not Analyzed	96.2	70.6	40.1	Not Analyzed
	Std. Deviation,%		2.7	4.7	4.0	
	% of Tests:					
	Within Specs., %		99.7	95.3	92.7	
	Above,%		0.0	0.0	7.3	
	Below,%		0.3	4.7	0.0	
Gravel Producer P (300 Tests)	Avg. Gradation,%		97.3	71.8	38.3	
	Std. Deviation,%		2.2	5.4	5.1	
	% of Tests:					
	Within Specs., %		100.0	91.7	91.7	
	Above,%		0.0	0.3	4.3	
	Below,%		0.0	8.0	4.0	
Gravel Producer R (300 Tests)	Avg. Gradation,%		98.0	70.4	38.9	
	Std. Deviation,%		2.2	4.5	4.1	
	% of Tests:					
	Within Specs., %		100.0	93.3	96.0	
	Above,%		0.0	0.0	3.0	
	Below,%		0.0	6.7	1.0	

TABLE 1 (cont.)
SUMMARY OF GRADATION TESTS ON ACCEPTED MATERIAL OF DIFFERENT AGGREGATE TYPES

Material and Producer	Item	Type 23A			
		Sieve Size and Spec. Limits, %			Loss by Washing 7-15 %
		1 in. 100	3/8 in. 60-85	No. 10 25-50	
Gravel Producer C (282 Tests)	Avg. Gradation,%		73.3	41.2	
	Std. Deviation,%		5.7	7.5	
	% of Tests:				
	Within Specs., %		98.6	90.0	
	Above,%		0.0	8.9	
	Below,%		1.4	1.1	
Gravel Producer J (300 Tests)	Avg. Gradation,%	Not Analyzed	74.8	46.3	Not Analyzed
	Std. Deviation,%		4.4	4.4	
	% of Tests:				
	Within Specs., %		99.7	86.0	
	Above,%		0.3	14.0	
	Below,%		0.0	0.0	
Gravel Producer L (300 Tests)	Avg. Gradation,%		78.2	43.5	
	Std. Deviation,%		4.9	6.1	
	% of Tests:				
	Within Specs., %		93.3	86.0	
	Above,%		6.7	14.0	
	Below, %		0.0	0.0	

Material and Producer	Item	Type 24A			
		Sieve Size and Spec. Limits, %			Loss by Washing 3-7 %
		1 in. 100	3/8 in. 60-85	No. 10 30-50	
Gravel Producer J (300 Tests)	Avg. Gradation,%		75.2	43.6	
	Std. Deviation,%		4.9	4.0	
	% of Tests:				
	Within Specs., %		99.4	95.7	
	Above,%		0.3	4.3	
	Below,%		0.3	0.0	
Gravel Producer K (300 Tests)	Avg. Gradation,%	Not Analyzed	74.5	44.3	Not Analyzed
	Std. Deviation,%		6.1	5.3	
	% of Tests:				
	Within Specs., %		98.0	90.3	
	Above,%		2.0	9.0	
	Below,%		0.0	0.7	
Gravel Producer M (300 Tests)	Avg. Gradation,%		68.6	41.4	
	Std. Deviation,%		5.2	5.5	
	% of Tests:				
	Within Specs., %		98.3	96.0	
	Above,%		0.0	2.3	
	Below, %		1.7	1.7	

Material and Producer	Item	Type 25A					
		Sieve Size and Spec. Limits, %					Loss by Washing 3 % max.
		5/8 in. 100	1/2 in. 90-100	3/8 in. 50-80	No. 4 10-25	No. 10 0-10	
Crushed Gravel Producer A (300 Tests)	Avg. Gradation,%		96.6	60.9	16.2	4.0	
	Std. Deviation,%		1.8	6.7	3.6	1.5	
	% of Tests:						
	Within Specs., %		99.7	98.0	99.4	100.0	
	Above,%		0.0	0.0	0.3	0.0	
	Below,%		0.3	2.0	0.3	0.0	
Stone Producer G (133 Tests)	Avg. Gradation,%	Not Analyzed	94.6	68.3	14.0	2.6	Not Analyzed
	Std. Deviation,%		2.2	6.2	3.3	1.1	
	% of Tests:						
	Within Specs., %		98.5	99.2	96.2	100.0	
	Above,%		0.0	0.0	0.0	0.0	
	Below,%		1.5	0.8	3.8	0.0	
Stone Producer Q (248 Tests)	Avg. Gradation,%		92.5	61.5	14.7	2.5	
	Std. Deviation,%		2.1	6.1	3.2	1.0	
	% of Tests:						
	Within Specs., %		97.6	99.2	98.8	100.0	
	Above,%		0.0	0.0	0.0	0.0	
	Below, %		2.4	0.8	1.2	0.0	

TABLE 1 (cont.)
SUMMARY OF GRADATION TESTS ON ACCEPTED MATERIAL OF DIFFERENT AGGREGATE TYPES

Material and Producer	Item	Type 26B					
		Sieve Size and Spec. Limits, %					Loss by Washing 3 % max.
		5/8 in. 100	1/2 in. 90-100	3/8 in. 60-85	No. 4 10-35	No. 10 0-10	
Gravel Producer A (157 Tests)	Avg. Gradation,%		97.7	70.9	17.0	2.1	
	Std. Deviation,%		1.6	7.7	4.5	1.3	
	% of Tests:						
	Within Specs., %		100.0	92.4	97.5	100.0	
	Above,%		0.0	0.6	0.0	0.0	
	Below,%		0.0	7.0	2.5	0.0	
Gravel Producer B (131 Tests)	Avg. Gradation,%	Not Analyzed	97.5	72.7	20.2	4.9	Not Analyzed
	Std. Deviation,%		1.4	8.2	5.9	2.5	
	% of Tests:						
	Within Specs., %		100.0	97.7	99.2	95.4	
	Above,%		0.0	0.0	0.8	4.6	
	Below,%		0.0	2.3	0.0	0.0	
Gravel Producer C (163 Tests)	Avg. Gradation,%		97.2	75.3	22.6	1.8	
	Std. Deviation,%		1.9	7.4	6.3	1.9	
	% of Tests:						
	Within Specs., %		100.0	90.1	98.2	100.0	
	Above,%		0.0	7.4	1.2	0.0	
	Below,%		0.0	2.5	0.6	0.0	

Material and Producer	Item	Type 2NS							
		Sieve Size and Spec. Limits, %							Loss by Washing 3 % max.
		3/8 in. 100	No. 4 95-100	No. 8 65-95	No. 16 35-75	No. 30 20-55	No. 50 10-30	No. 100 0-10	
Sand Producer A (300 Tests)	Avg. Gradation,%			86.4	69.3		18.0	2.6	
	Std. Deviation,%			3.0	3.2		3.5	1.2	
	% of Tests:								
	Within Specs., %			100.0	99.0		100.0	100.0	
	Above,%			0.0	1.0		0.0	0.0	
	Below,%			0.0	0.0		0.0	0.0	
Sand Producer B (300 Tests)	Avg. Gradation,%			86.3	61.9		16.1	3.2	
	Std. Deviation,%			4.3	3.5		3.2	0.9	
	% of Tests:								
	Within Specs., %			100.0	100.0		99.0	100.0	
	Above,%			0.0	0.0		0.0	0.0	
	Below,%			0.0	0.0		1.0	0.0	
Sand Producer C (300 Tests)	Avg. Gradation,%	Not Analyzed	Not Analyzed	80.6	55.9	Not Analyzed	13.5	2.7	Not Analyzed
	Std. Deviation,%			3.4	6.3		3.4	0.9	
	% of Tests:								
	Within Specs., %			100.0	100.0		98.7	100.0	
	Above,%			0.0	0.0		0.3	0.0	
	Below,%			0.0	0.0		1.0	0.0	
Sand Producer D (300 Tests)	Avg. Gradation,%			80.7	66.1		13.3	3.2	
	Std. Deviation,%			2.8	3.2		2.8	1.2	
	% of Tests:								
	Within Specs., %			100.0	100.0		98.3	100.0	
	Above,%			0.0	0.0		0.0	0.0	
	Below,%			0.0	0.0		1.7	0.0	
Sand Producer E (300 Tests)	Avg. Gradation,%			89.6	69.4		18.6	3.3	
	Std. Deviation,%			2.5	3.4		2.9	0.8	
	% of Tests:								
	Within Specs., %			100.0	99.0		100.0	100.0	
	Above,%			0.0	1.0		0.0	0.0	
	Below,%			0.0	0.0		0.0	0.0	

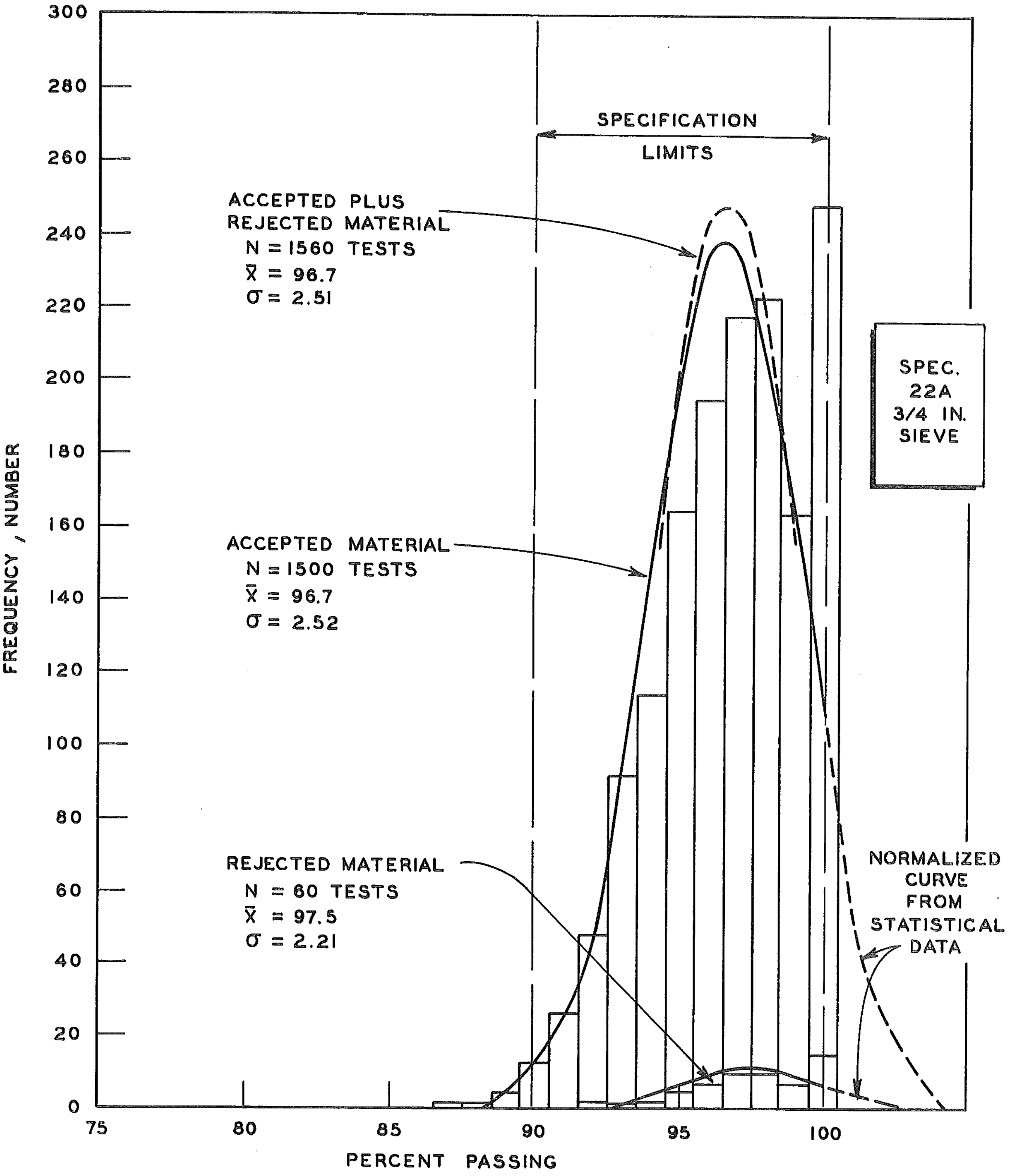


Figure 1. Distribution of percent passing 3/4-in. sieve for accepted material, rejected material, and the combination for 22A aggregate.

TABLE 2
SUMMARY OF ACCEPTED AND REJECTED 22A AGGREGATE BATCHES

Project	Number of Test Results		
	Total	Accepted	Rejected
1	311	307	4
2	319	306	13
3	338	319	19
4	313	313	0
5	245	210	35
6	136	136	0
7	79	79	0
8	207	202	5
9	169	166	3
10	461	454	7
11	246	228	18
	2824	2720	104

Figures 1, 2, and 3 show the effect of rejected material on frequency distributions of percents passing the 3/4- and 3/8-in., and No. 10 sieves, respectively.

Although tests were studied for 104 lots of 150 tons each of rejected material (out of 2,824 lots), the plotting of the distribution of percents passing certain sieves was confined to 60 samples selected at random from the 104. This reduction was made to maintain the approximate proportions of accepted and rejected material. The three graphs show that inclusion of rejected material in the statistical analysis at the rate of approximately 3.7 percent, as encountered for 22A aggregate, would not significantly alter the averages or standard deviations given in Table 1. Thus, for at least this one case, the overall distribution is not significantly changed by this inclusion of rejected material.

Figure 4 shows the distribution of non-conforming tests by type of aggregate, producer, and various sieve sizes for gravel, stone, and slag. The non-conforming tests shown are those only within the population of accepted material— that is, the same data given in Table 1. For 4A aggregate and 2NS sand, the distribution of the percents passing is indicated for various

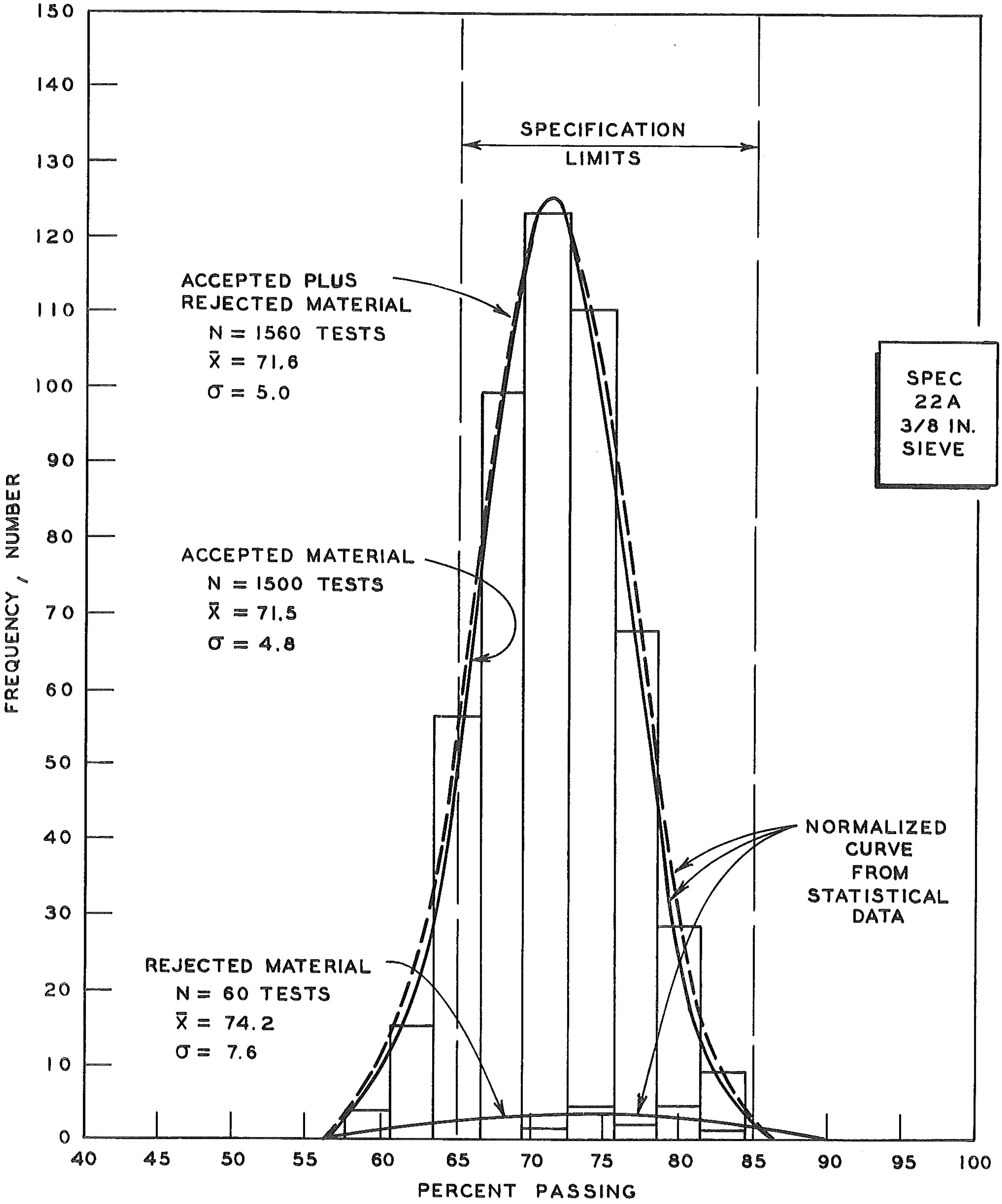


Figure 2. Distribution of percent passing 3/8-in. sieve for accepted material, rejected material, and the combination for 22A aggregate.

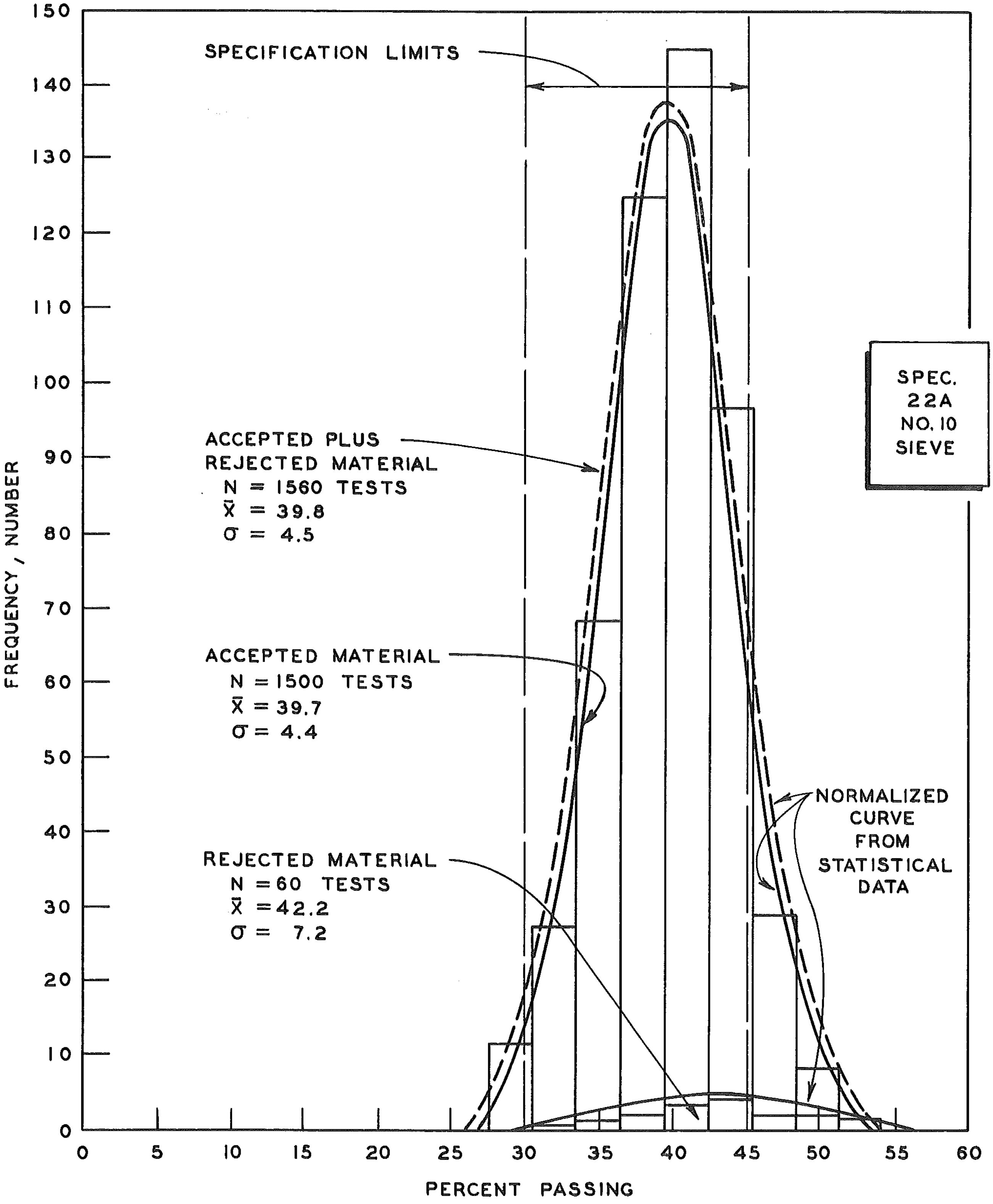


Figure 3. Distribution of percent passing No. 10 sieve for accepted material, rejected material, and the combination for 22A aggregate.

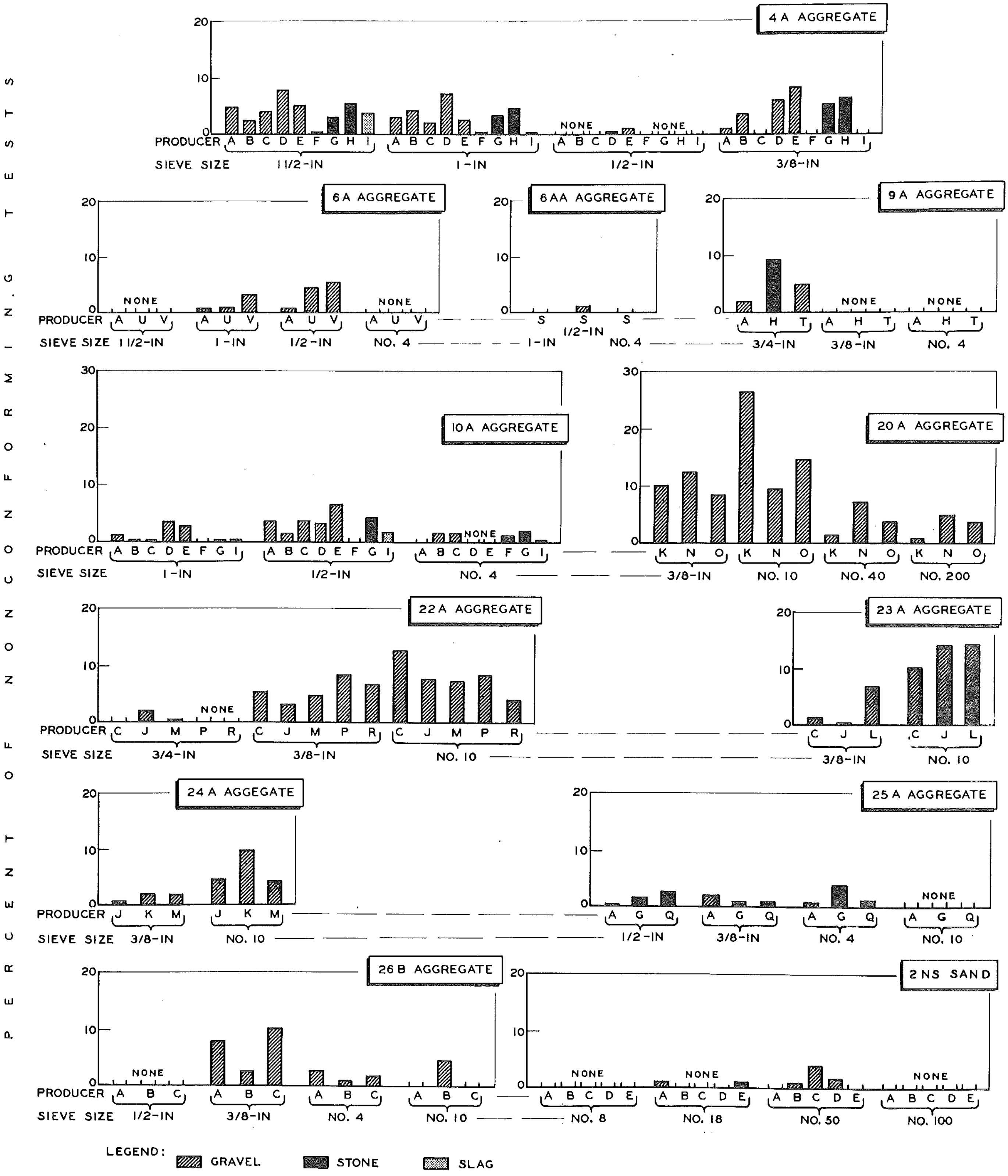


Figure 4. Variability of fractions defective (nonconforming tests) among aggregate producers for materials passing various sieve sizes.

sieve sizes. For 4A aggregate (Fig. 5), the distribution of percent passing the 1/2- and 3/8-in. sieves tends to be close to the lower specification limit. For 2NS sand (Fig. 5), the distribution of percents passing the Nos. 8 and 16 sieves bunch toward the upper specification limit, for the No. 50 it is pretty well centered, but for the No. 100 sieve the results definitely bunch toward the lower specification limit of zero. Table 1 illustrates the inherent distribution characteristics within the specification limits for the other type aggregates.

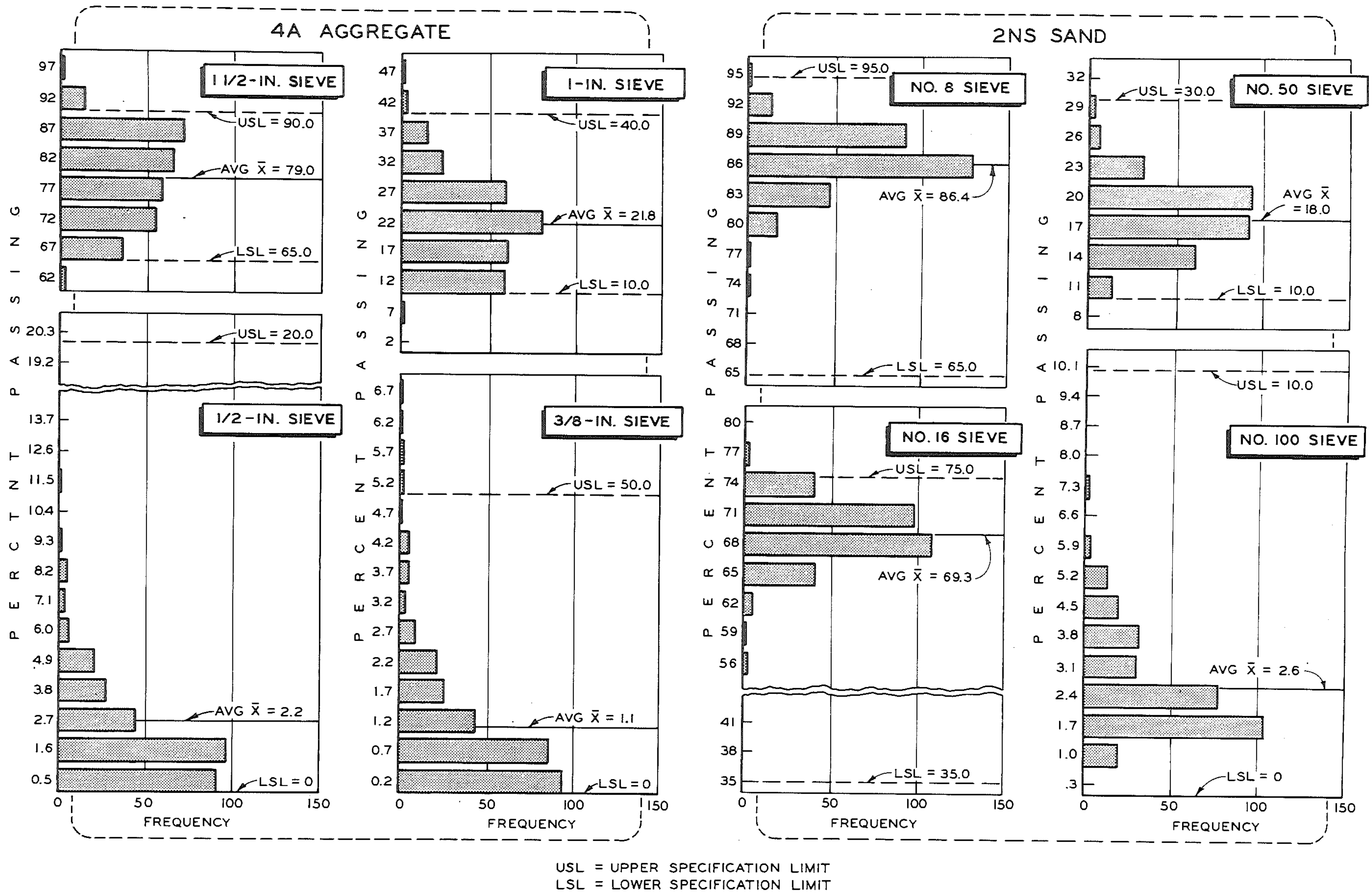


Figure 5. Histograms for 4A aggregate and 2NS sand.

FINENESS MODULUS

Fineness modulus of either fine or coarse aggregate is a measure of the degree of uniformity of its grading. It is used as an index of coarseness or fineness—the higher the modulus, the coarser the aggregate. In Michigan, fineness modulus is computed only for sand, by adding the cumulative percentages by weight of the material retained on the 3/8-in., Nos. 4, 8, 16, 30, 50, and 100 sieves, and dividing the sum by 100. The base fineness modulus for each pit is determined from the average of a continued period of production of acceptable fine aggregate. The specifications for gradation of fine aggregate (2NS sand) require control of grading so that the fineness modulus of representative samples should not vary more than ±0.20 from the average fineness modulus of all samples previously taken from the same source.

From daily records of aggregate field testing, covering the period of 1962 through 1963, ten projects were sampled at random to study relative variability of the fineness modulus of 2NS sand. Approximately 20 samples of 2NS gradings were randomly selected from each of the ten to compute the modulus according to specifications. The results of these computations are summarized in Table 3.

In addition, Figure 6 shows several daily fineness modulus averages on natural (2NS) sand, plotted in chronological order. These samples were all taken from one belt. Starting with Report No. 9, a base modulus of 3.06 was established using the running average of the first six tests, with limits of ±0.20 (indicated by the darker shading at center). The average of the entire group is 2.98. However, a clear downward trend is evident after Report No. 13 and again after Report No. 14, with the material becoming progressively finer through Report No. 23; production was stopped for the winter after Report No. 25. Of interest here is the wide variation in daily results from a given pit.

Corresponding information from a different pit is shown in Figure 7 for samples taken from one stockpile. Considerably more uniformity is evident here. A base of 2.81 was established from results of the first 13 reports, which were taken prior to the tests graphed in Figure 7.

TABLE 3
SUMMARY OF TEST RESULTS ON FINENESS MODULUS FOR GRADATION OF 2NS SAND

Project No. And Date	Number of Tests	Fineness Modulus	
		Average	Standard Deviation
47014, C6 (11-15-61 to 9-6-62)	20	2.85	0.11
33084, C5 (1-25-62 to 8-29-62)	20	2.87	0.11
63081E, C10 (4-20-62 to 10-11-62)	13	2.79	0.10
81074, C1 (5-23-62 to 10-18-62)	24	2.94	0.11
63173, C4 (6-27-62 to 7-18-62)	20	2.83	0.06
47013, C8 (7-17-62 to 10-1-62)	21	3.26	0.32
49025, C18 (10-5-62 to 8-19-63)	20	2.89	0.27
47065, C1 (10-13-62 to 10-18-62)	20	3.31	0.12
41027A, C (10-18-62 to 6-3-63)	20	2.88	0.08
63101A, C1 (10-19-62 to 7-2-63)	22	3.44	0.09

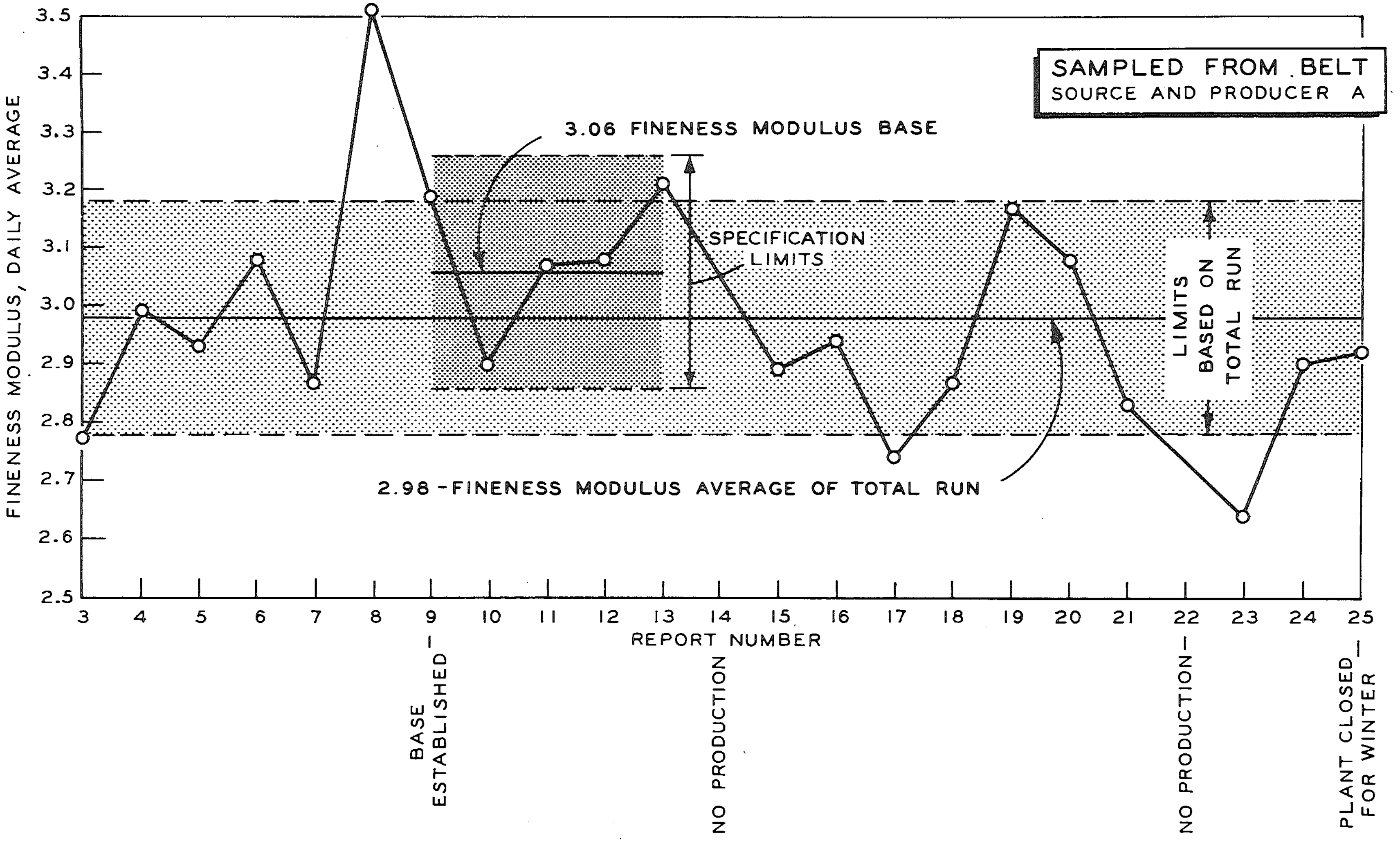


Figure 6. Fineness modulus for natural (2NS) sand sampled from belt.

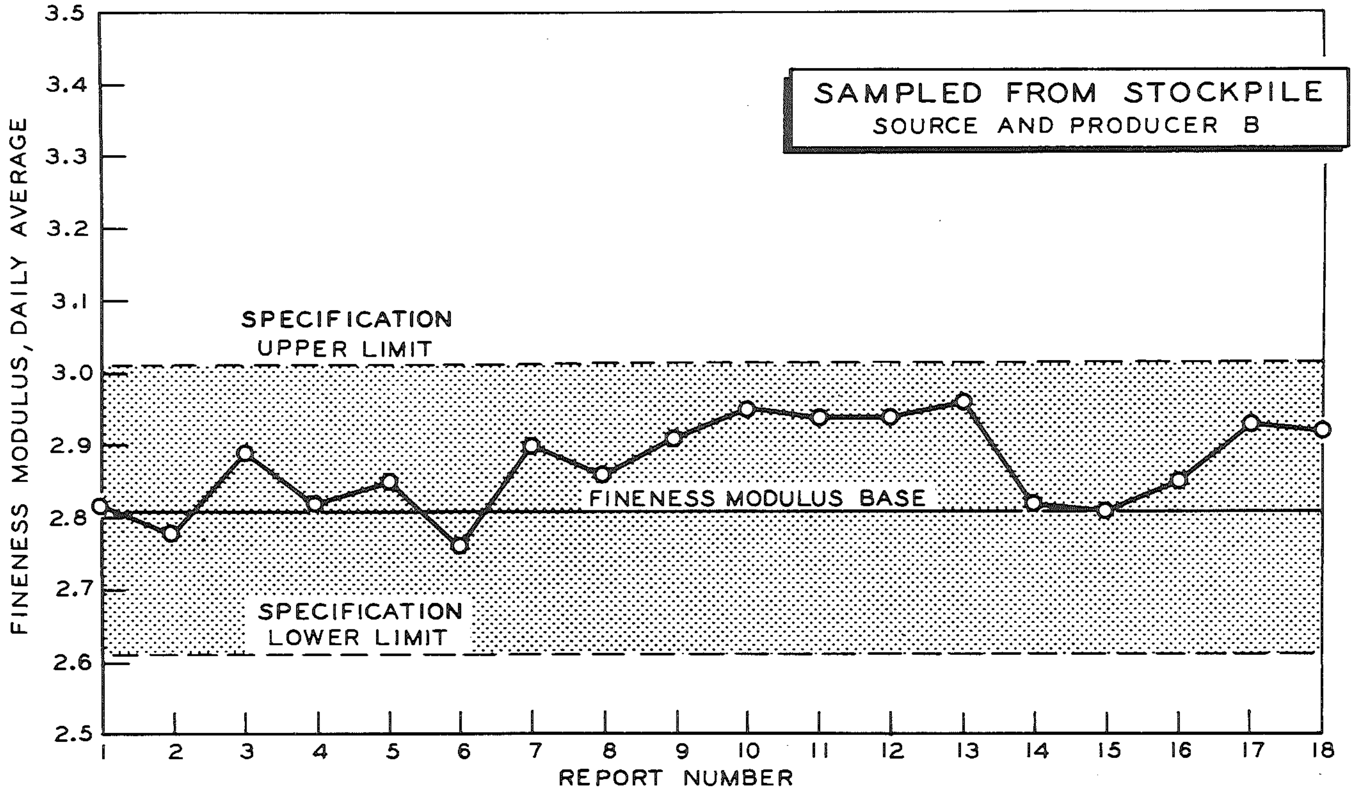


Figure 7. Fineness modulus for natural (2NS) sand sampled from stockpile.

FLEXURAL STRENGTH OF CONCRETE BEAMS

Flexural strength of concrete beams was studied from field records of all 1959-1961 Michigan projects using Grade A concrete for both pavement and bridge construction. Hundreds of field tests found in project files were condensed on special form sheets and punched on data cards for computer processing. Frequency distributions were grouped and analyzed for both road and bridge concrete beam specimens 7, 14, and 28 days old. The analysis involved computing the mean value or sample average $\overline{X}$, the sampling variability or standard deviation S, and the relative variability or coefficient of variation V. The relative variability here was computed by the expression:

$$V = 100 \frac{S}{\overline{X}} \qquad (1)$$

which is useful for indicating the degree of uniformity in flexural strength tests on concrete beam specimens. From these computations the percentage of tests below the specified minimum limits (fraction defective) was estimated. The estimated values were checked against actual defective values found from the cumulative frequency distributions.

The results covering approximately three years of construction work under varying weather conditions are summarized in Table 4 for both pavement and bridge projects. Variations in strength within single beams were computed by multiplying the average range of several groups of two breaks per beam by a constant of 0.8865, as given in the literature. Results are summarized in Table 5 for both pavement and bridge projects. For this table, within-test relative variation V_1 was computed as follows:

$$V_1 = 100 \frac{S}{\overline{X}} \text{ and } S_1 = \frac{\overline{R}}{d_2} \qquad (2)$$

where: S_1 = within-test standard deviation

$\frac{1}{d_2}$ = 0.8865 from statistical tables

$\overline{R}$ = average range of groups of two breaks per beam

$\overline{X}$ = average strength.

According to present specifications for flexural strength (modulus-of-rupture), the steel molds in which the beams are cast are nominally 6 by 6 by 36 in. A set of four beams is made on alternate days when the pavement is placed; for concrete structures, one set of beams is made from the first concrete pour at the job site, and a set for each succeeding 200 cu yd

TABLE 4
SUMMARY OF FLEXURAL STRENGTH TESTS ON 1959-61 CONCRETE BEAM SPECIMENS

			Item	Individual Values			Average Values of Two Breaks per Beams			Minimum Values Within Two Breaks per Beam		
				Spring	Summer	Fall	Spring	Summer	Fall	Spring	Summer	Fall
Pavement Concrete	Cement Factor = 5.5 sacks per cu yd	7-day Tests	Number of tests	713	1904	701	354	942	345	354	942	345
			Average strength, psi	654.9	652.7	644.8	655.3	653.3	645.3	632.7	631.9	619.6
			Relative variation, percent	12.9	13.9	15.5	11.9	13.1	14.7	12.0	13.2	15.3
			Percent less than 550 psi	8.0	9.0	16.5	5.5	7.0	15.0	11.0	11.0	22.0
		14-day Tests	Number of tests	601	1604	596	300	792	294	300	792	294
			Average strength, psi	749.3	750.4	740.6	749.6	751.2	741.7	726.9	729.3	714.6
			Relative variation, percent	12.3	18.1	15.5	11.6	17.7	14.7	12.2	17.8	15.5
			Percent less than 600 psi	3.0	5.5	11.5	2.5	4.5	10.0	6.5	9.0	15.0

			Item	Individual Values				Average Values of Two Breaks per Beam				Minimum Values Within Two Breaks per Beam			
				Spring	Summer	Fall	Winter	Spring	Summer	Fall	Winter	Spring	Summer	Fall	Winter
Bridge Concrete	Cement Factor = 5.9 sacks per cu yd	7-Day Tests	Number of tests	635	778	474	181	317	388	236	89	317	388	236	89
			Average strength, psi	654.1	661.7	620.7	590.9	654.3	661.7	620.1	593.0	633.4	639.5	600.5	565.7
			Relative variation, percent	13.8	15.0	14.2	25.2	13.1	14.3	13.2	24.1	13.4	14.2	13.5	25.0
			Percent less than 550 psi	11	10	19	34	11	9	21	28	16	14	27	42
		28-Day Tests	Number of tests	608	742	467	156	304	371	233	78	304	371	233	78
			Average strength, psi	814.2	802.0	777.7	726.3	814.2	802.0	777.9	727.0	793.4	778.1	754.6	700.3
			Relative variation, percent	13.3	13.8	13.6	17.2	12.8	13.2	13.0	16.4	13.0	13.5	13.4	16.6
			Percent less than 650 psi	5	6	10	20	5	5	9	19	8	9	13	25
	Cement Factor = 5.5 sacks per cu yd	7-Day Tests	Number of tests	294	459	262	153	147	221	130	76	147	221	130	76
			Average strength, psi	614.8	626.8	591.9	571.4	615.1	628.8	591.8	570.4	592.9	607.1	572.4	549.3
			Relative variation, percent	15.2	13.8	15.9	23.5	14.3	13.1	15.3	23.0	14.1	13.5	15.6	23.3
			Percent less than 550 psi	23	18	29	37	21	13	26	39	28	22	34	50
		28-Day Tests	Number of tests	283	462	252	174	141	231	126	87	141	231	126	87
			Average strength, psi	795.9	762.5	756.6	709.2	796.6	762.5	756.6	709.2	771.6	741.1	734.2	684.6
			Relative variation, percent	13.6	13.6	13.8	16.4	12.9	13.1	13.2	15.6	13.3	13.3	13.5	16.2
			Percent less than 650 psi	9	15	16	28	6	13	15	25	12	19	19	33

NOTE: There are normally two breaks per beam. In this survey, an occasional report gave only one break, or only one was legible, resulting in slight discrepancies between columns. Results are presented in this way so that the reader may select the information of interest.

TABLE 5
WITHIN-TEST RELATIVE VARIATIONS FOR FLEXURAL STRENGTH
1959-61 Concrete Beam Specimens

			Item	Spring	Summer	Fall	Winter
Pavement Concrete	Cement Factor = 5.5 sacks per cu yd	7-Day Tests	Number of tests	100	100	100	---
			Avg. strength, psi	655.3	653.3	645.3	---
			Average range, psi	42.5	38.4	51.8	---
			Relative variation, percent	5.7	5.2	7.1	---
		14-Day Tests	Number of tests	100	100	100	---
			Average strength, psi	749.6	751.2	741.7	---
			Average range, psi	44.0	46.5	58.4	---
			Relative variation, percent	5.2	5.5	7.0	---
Bridge Concrete	Cement Factor = 5.9 sacks per cu yd	7-Day Tests	Number of tests	100	100	100	89
			Average strength, psi	654.3	661.7	620.1	593.0
			Average range, psi	39.9	39.9	39.8	54.3
			Relative variation, percent	5.4	5.3	5.7	8.1
		28-Day Tests	Number of tests	100	100	100	100
			Average strength, psi	814.2	802.0	777.9	727.0
			Average range, psi	37.4	45.9	39.7	51.4
			Relative variation, percent	4.1	5.1	4.5	6.3
	Cement Factor = 5.5 sacks per cu yd	7-Day Tests	Number of tests	100	100	100	76
			Average strength, psi	615.1	628.8	591.8	570.4
			Average range, psi	46.0	39.8	42.7	41.0
			Relative variation, percent	6.6	5.6	6.4	6.4
		28-Day Tests	Number of tests	100	100	100	100
			Average strength, psi	796.6	762.5	756.6	709.2
			Average range, psi	49.0	47.4	46.1	48.6
			Relative variation, percent	5.5	5.5	5.4	6.1

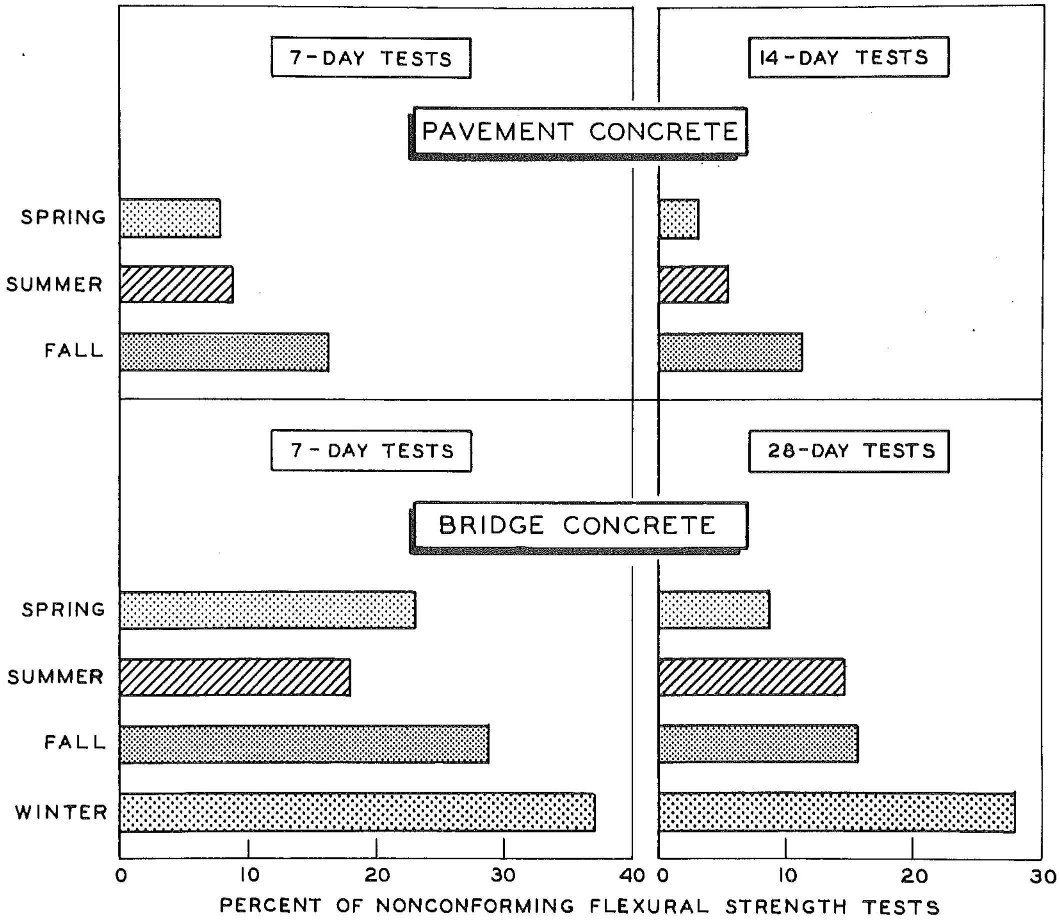


Figure 8. Flexural strength of concrete beams, reflecting influence of seasonal testing conditions on field results.

poured. For pavements, two beams are tested at 7 and 14 days, and for bridges, two beams are tested at 7 and 28 days. Michigan uses the cantilever type of loading, with two breaks per beam.

The minimum values (lower acceptance limits) required for flexural strength when concrete specimens are properly cured under moist conditions at 60 and 80 F are 550 psi at 7 days, 600 psi at 14 days, and 650 psi at 28 days for cement factors of 5.5 or 5.9 sacks per cu yd of concrete.

Two facts were noted in connection with the resulting frequency distributions. First, the grouped data from both pavements and bridges followed closely symmetrical distributions. Second, the grouped data reflected the influence of seasonal testing conditions. This can be explained in terms of fluctuations of relative variation for individual and average values of two breaks per beam as shown in Table 4. Values for percent of nonconforming tests, for example, were significantly higher in fall than in spring for pavement projects, and higher in winter than in spring for bridge projects, as shown in Figure 8. Some generalizations may be made from Tables 4 and 5 regarding effects of the flexural strength average, the relative variation, and the specified lower acceptance limits required in flexural strength tests:

1) Strength tests for pavement concrete in seasons other than spring, with approximately the same average value but greater relative variation than those recorded in spring, show a greater proportion of values falling below the lower limit.

2) Fluctuations in relative variations within single beams reflect the influence of seasonal curing conditions.

THICKNESS OF CONCRETE PAVEMENTS

The study of thickness of portland cement concrete pavements was based on 15 projects sampled at random from field records covering the period of 1959-1961. From these data, 656 tests were grouped and analyzed to estimate standard statistical parameters as well as the percentage of tests outside the specified limits. Specifications require that thickness of the completed pavement be controlled so that the average depth of each concrete core should not be more than 1/2-in. under the design thickness. One core is taken at random from each 1,000-ft length of pour in accord with plan stationing. The depth of each core is found using a depth-gage from the center of the upper end of the specimen and from eight other points equally spaced around the center. The individual measurements are recorded to the nearest 0.05 in., and the average of these nine measurements expressed to the nearest 0.10 in. is considered the core depth.

The frequency distribution of measured pavement thickness as obtained from the average depth of each 656 concrete cores is shown in Figure 9. While the nominal pavement thickness is 9 in., the average depth measured was 9.2 in. and one standard deviation was ±0.28 in. The distribution is quite symmetrical and approximates a normal distribution. Only two cores out of 656 were below the minimum specification depth of 8.5 in., or approximately 0.3 percent.

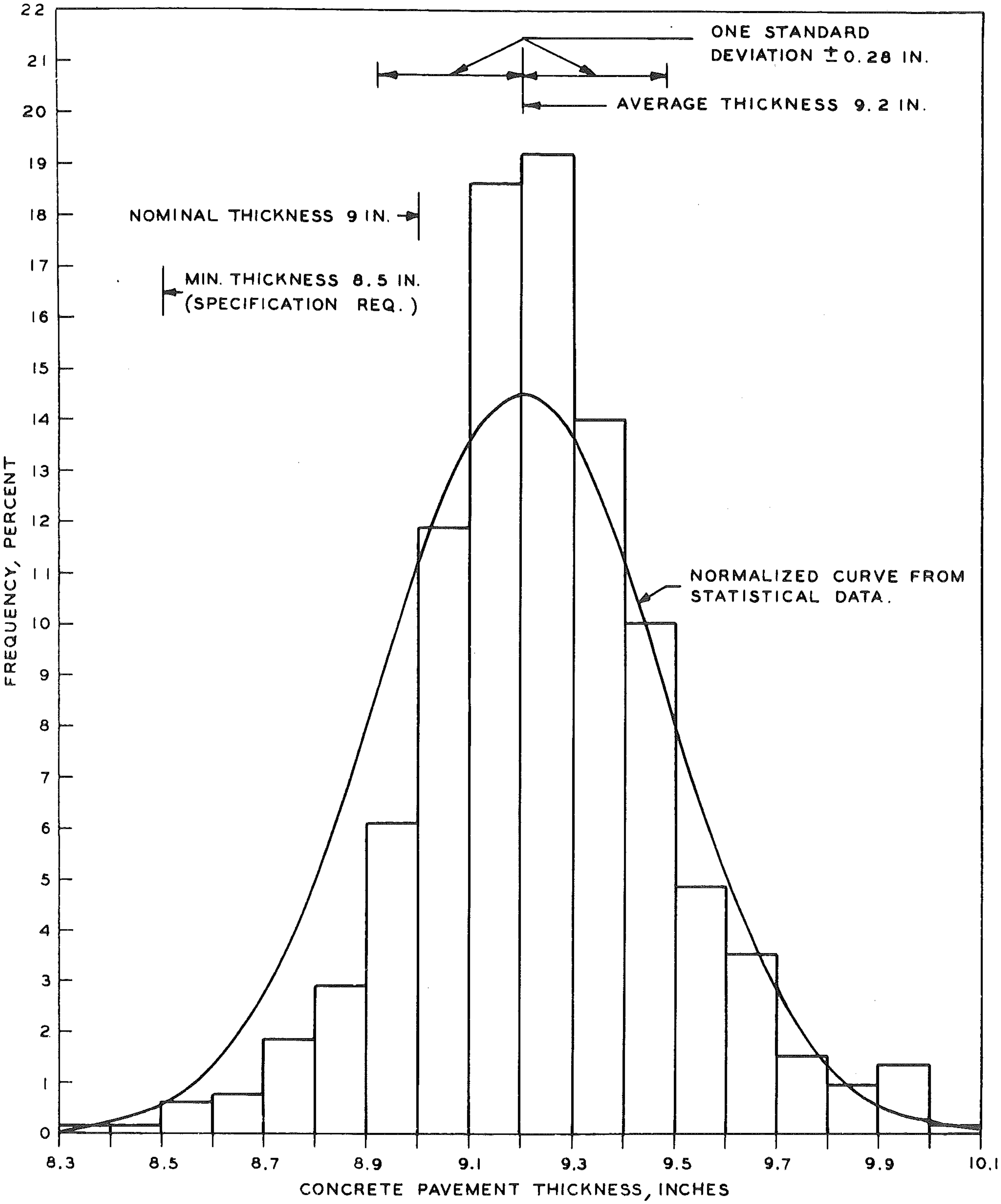


Figure 9. Frequency distribution of concrete pavement thickness: 1959-61.

AIR CONTENT OF FRESH PAVEMENT CONCRETE

Analysis of field data dealing with air content of fresh pavement concrete was based on 60 projects sampled at random from files covering the period from 1959-1961. From the selected projects 4,065 tests were grouped and analyzed to estimate important characteristics of the data, such as the degree of normality, average, and standard deviation of the test results. Michigan requires an average air content between 4 and 7 percent for concrete pavement. Periodic tests for air content of freshly mixed concrete are made at the job site by the standard pressure meter method.

As shown in Figure 10, the frequency distribution shows acceptable normality for this condition of symmetrical control limits, and the average of 5.65 percent is very close to the midpoint of the specification (5.5 percent). Slightly over 90 percent of the tests fall within the specification limit, with 2.9 percent below and 6.3 percent above the lower and upper limits, respectively.

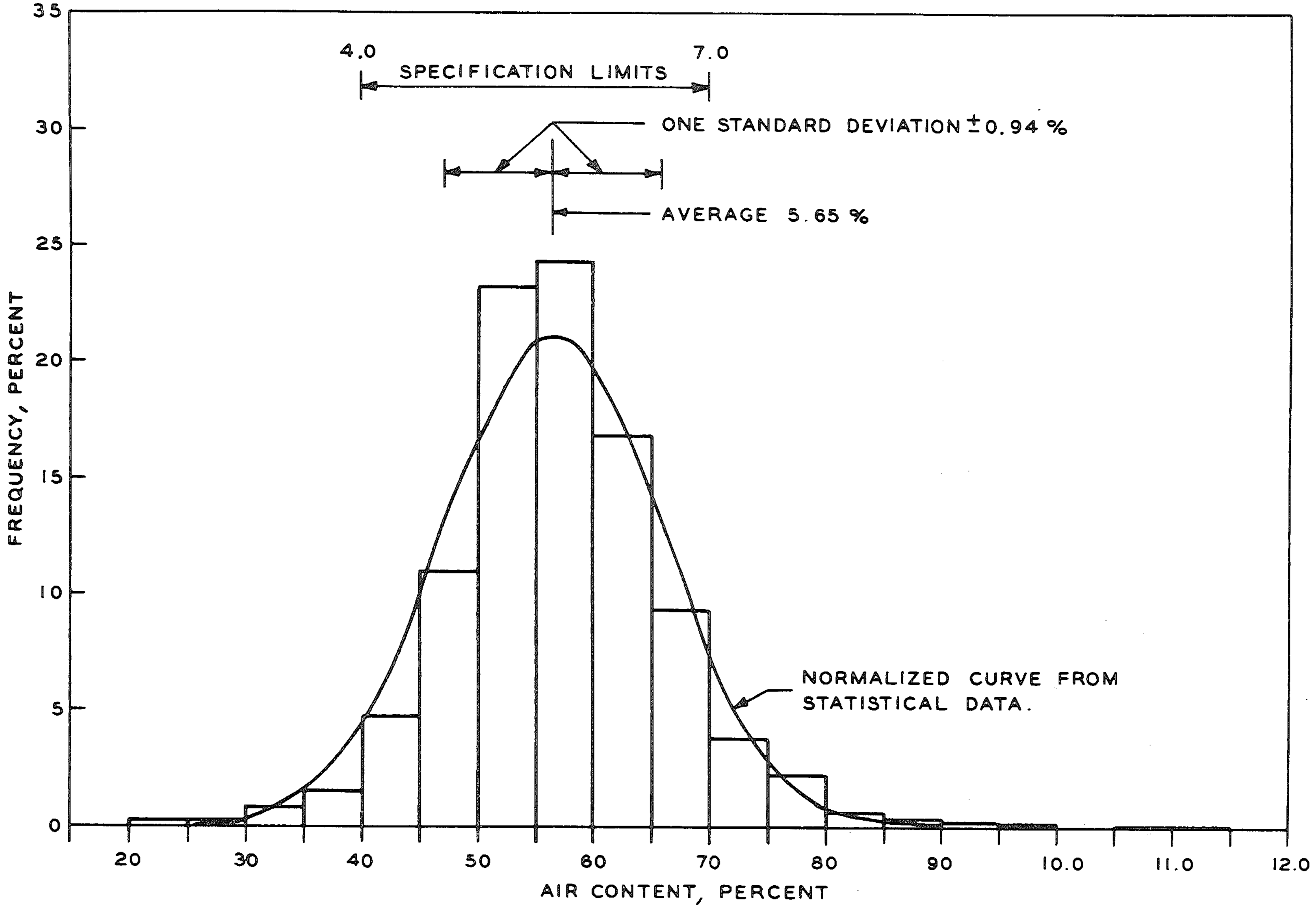


Figure 10. Frequency distribution of concrete pavement air content: 1959-61.

MOISTURE AND DENSITY OF AGGREGATE BASE COURSE AND SAND SUBBASE

Field data dealing with moisture content and density of subbase for pavements were analyzed, covering the 1961 and 1962 construction periods on the basis of 132 projects sampled at random from the files. In Figure 11, frequency distributions of moisture content for aggregate base course and sand subbase are shown. For the former, the average moisture content is 5.1 and for the latter 6.3 percent, while standard deviations are ±1.6 and ±2.6 percent, respectively. In both cases, a slightly unsymmetrical distribution may be noted with the long tail always at the higher end of the moisture content.

For aggregate base course and sand subbase, the Michigan Cone Test is used to obtain the maximum unit weight of the material and the Rainhart test is used routinely to determine actual density of the material as placed and compacted on the grade. Department specifications require a minimum of 100 percent of the maximum unit weight for aggregate base course and 95 percent for sand subbase. Routine inspection procedure calls for testing at locations giving evidence of being the least compacted. If the test indicates that the compaction does not meet the particular specification requirement (95 or 100 percent of the maximum unit weight) then further compaction of the area is required and a second test is made.

It should be noted that the frequency distributions shown in Figure 12 for aggregate base course and sand subbase are only for final or accepted tests— that is, tests meeting the specification and representing accepted compacted areas. The average density values for 1961 and 1962 construction seasons were nearly identical, being 102.2 and 102.1 percent, respectively, for 1961 and 1962 aggregate base course, and 98.2 and 98.4 percent for sand subbase for 1961 and 1962. Standard deviations for sand subbase were nearly identical in 1961 and 1962, being 2.16 and 2.14 percent. However, the standard deviation for aggregate base course increased from 2.05 percent in 1961 to 2.62 percent in 1962. Both materials and each year of construction presented frequency distributions with the long tail at the high end of the compaction range.

The preceding data were based on accepted tests only and, therefore, for a more complete picture of the total density distributions likely to be experienced, some additional density data were analyzed from a 1965 construction project near Lansing. On this project, 502 initial density measurements were made on the clay embankment, with 143 being below the 95 percent density requirement and 359 tests acceptable (95 percent or greater).

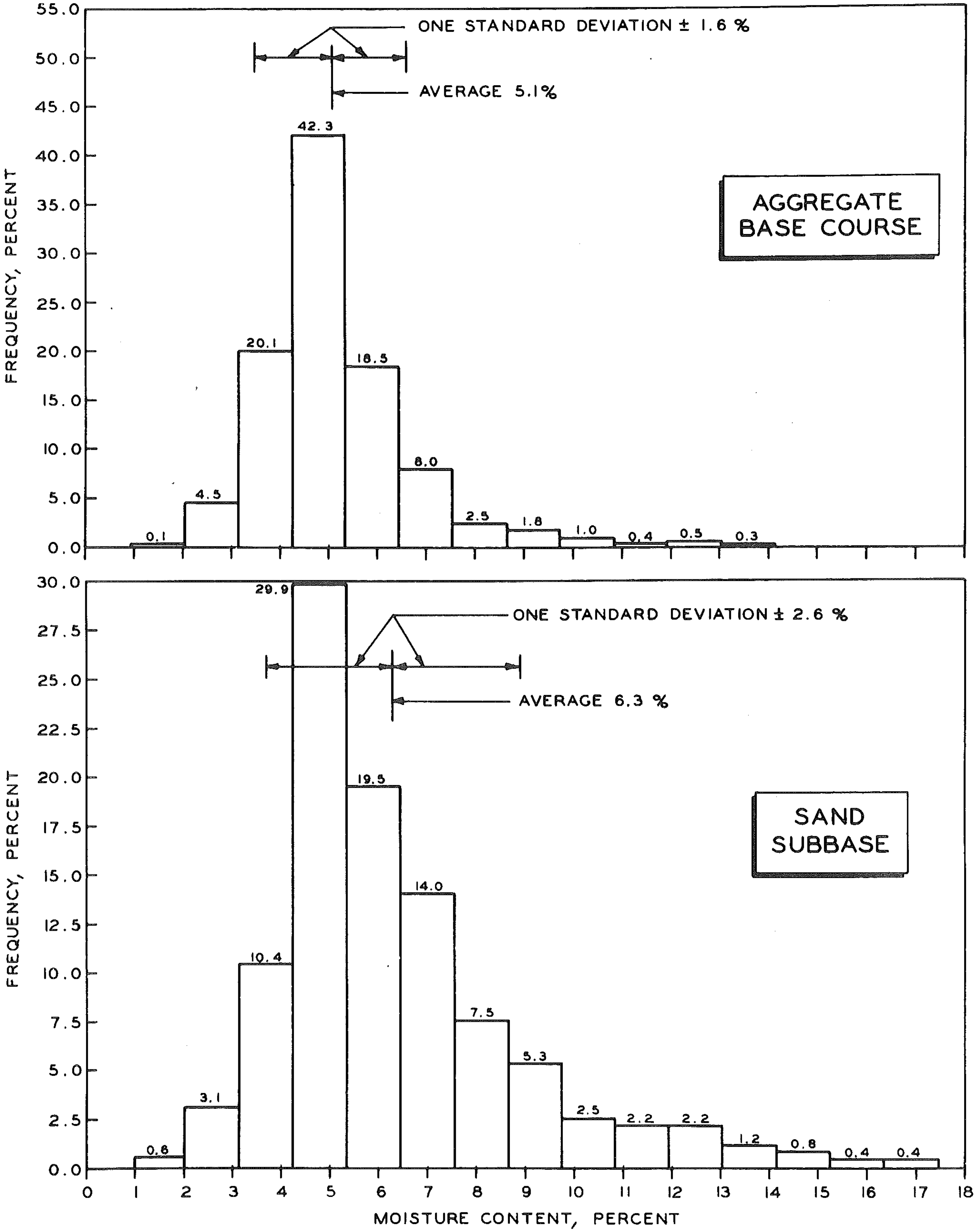


Figure 11. Frequency distributions of moisture content for aggregate base course (802 tests) and sand subbase (800 tests).

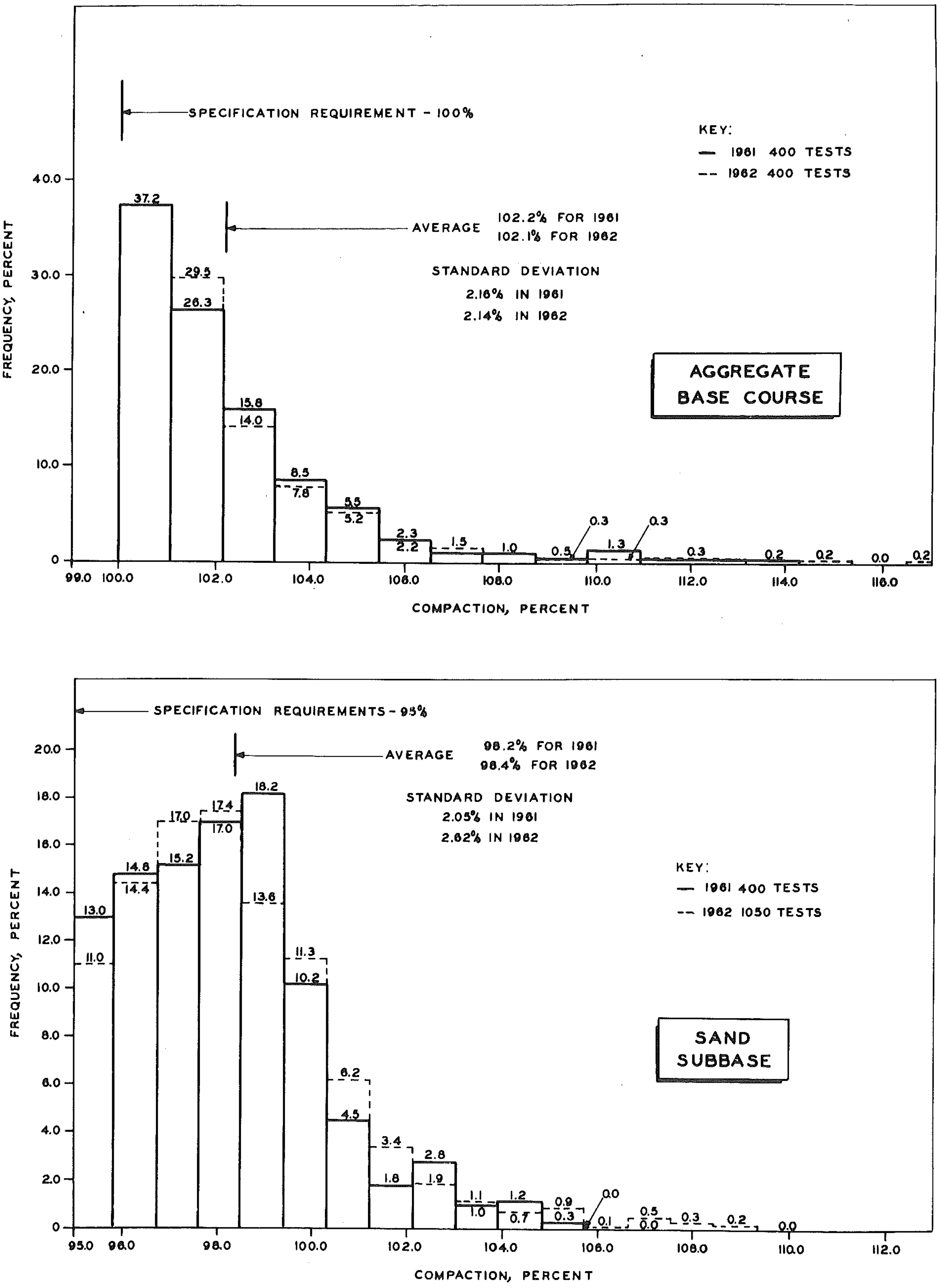
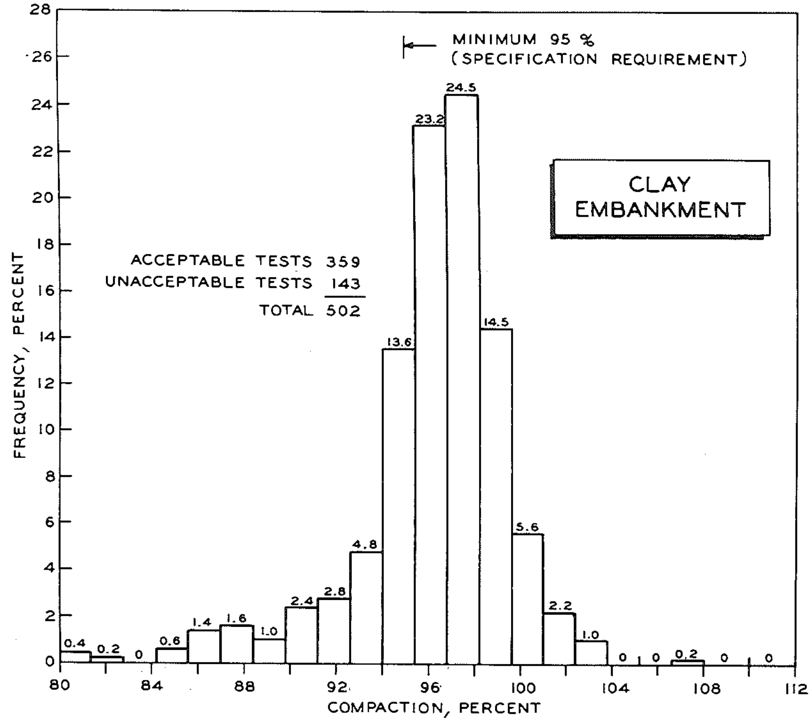
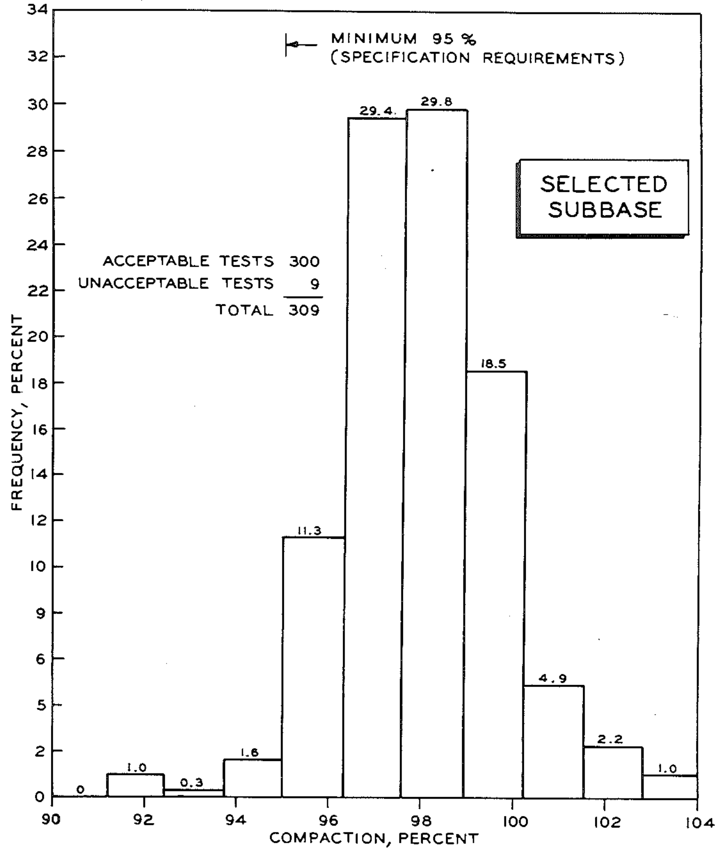


Figure 12. Frequency distributions of percent compaction for aggregate base course and sand subbase, for final or accepted tests only.





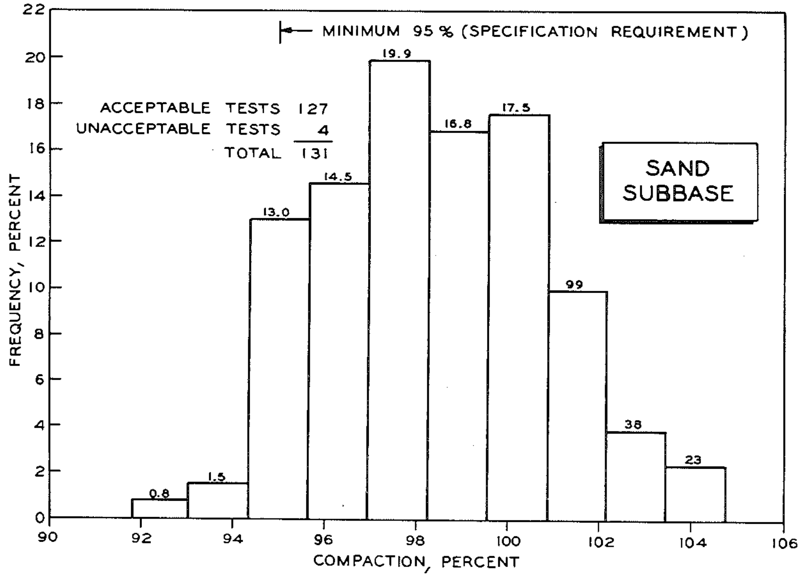


Figure 13. Frequency distributions of density test results for three soils; graphs exclude all retests, but include all initial failing tests.

This density distribution is illustrated in Figure 13. For all grade areas represented by unacceptable test results, additional compaction was required until further testing indicated that the grade was compacted to the minimum 95 percent of maximum unit weight. Figure 13 also illustrates similar density distributions of initial test results for selected subbase and sand subbase. For selected subbase, 2.9 percent were initially unacceptable, and for sand subbase 3.05 percent.

On this same project some random density tests were made of the selected subbase and sand subbase on grades already accepted on the basis of standard sampling and testing procedures. Figure 14 illustrates that the density distribution for accepted selected subbase varied from 90.8 to 103.8 percent of maximum unit weight, with approximately 20 percent of the random tests falling below the minimum requirement of 95-percent compaction. A similar distribution of random tests from previously accepted sand subbase grade is also illustrated in Figure 14, where approximately 34 percent of the test results were below the specification requirement of 95 percent of the maximum unit weight.

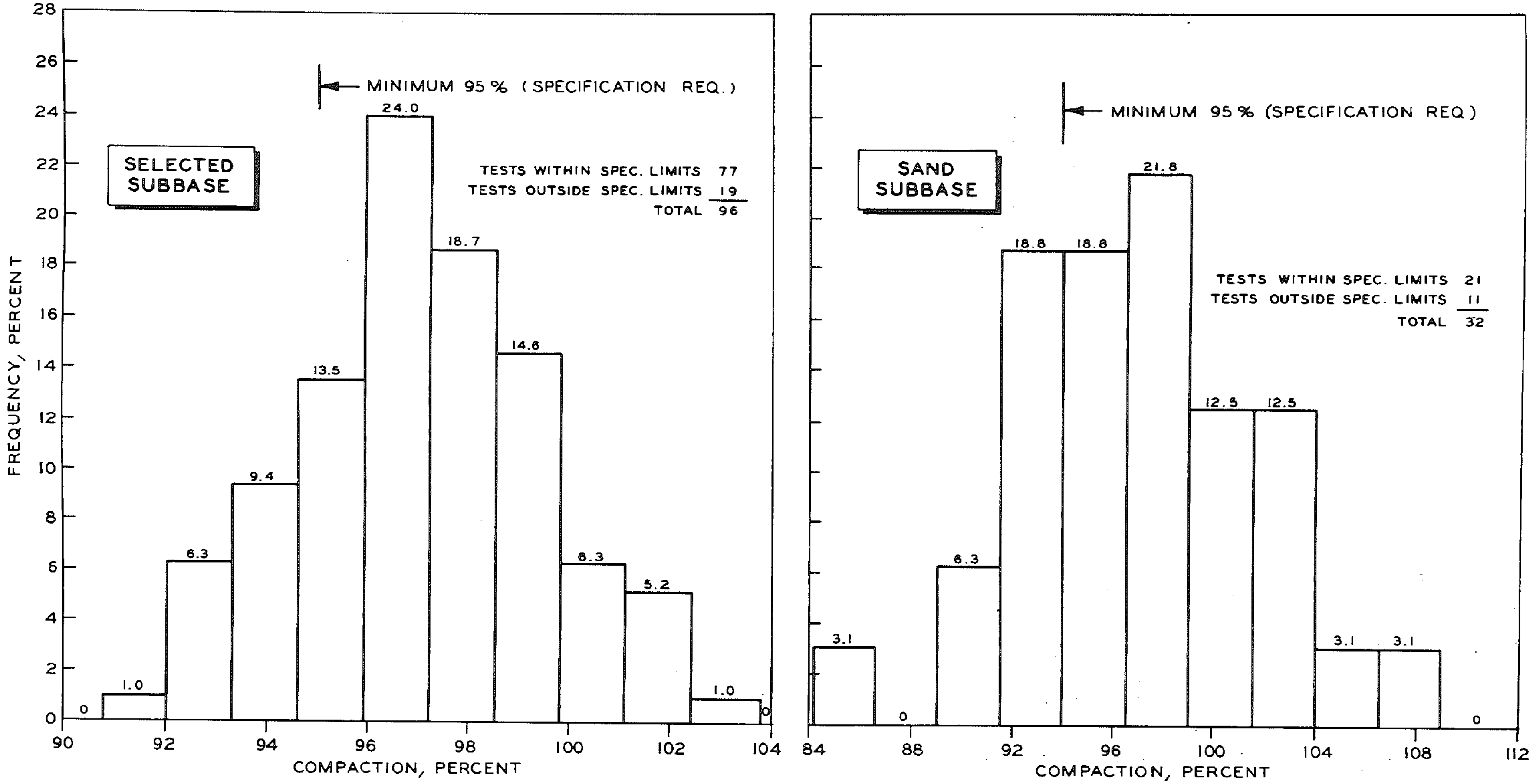


Figure 14. Frequency distributions of density test results for accepted aggregate base course and sand subbase.

AGGREGATE GRADATION— FIELD EXPERIMENT

The field records survey for aggregate gradation indicated that sampling and testing procedures now in use for acceptance or rejection of such materials allow many possibilities for introduction of deviations.

A field experiment was carefully planned to determine what part aggregate inspectors, screening sieves, and sampling methods play in the uniformity of gradation results. This required an experimental design which may be discussed in terms of the following mathematical equation:

$$X = G + I + M + S + IM + IS + MS + IMS + E \tag{3}$$

where:

X = an individual test
G = the overall mean
I, M, S = effects due to inspectors, sampling methods, and screening kits, respectively.
IM, IS, MS, IMS = effects due to various interactions.
E = random error

Based on this model and other algebraic identities the following variances may be determined:

V_t = overall variance of the gradation results
V_1 = variance attributable to different aggregate inspectors
V_2 = variance attributable to different screening sieves
V_3 = variance attributable to different sampling procedures
V_4 = variance attributable to inherent material and experimental deviations.

In planning the experimental work, consideration was given to the availability of manpower and testing equipment, and the type and location of the aggregate materials being produced, thus largely limiting the experiment to a fixed statistical model.

The experiment was expected to indicate:

1) Whether aggregate inspectors require further training in sampling and testing of aggregate materials.
2) Whether testing equipment requires periodic calibration or more careful maintenance.
3) Whether improved precision is feasible in gradation analysis.

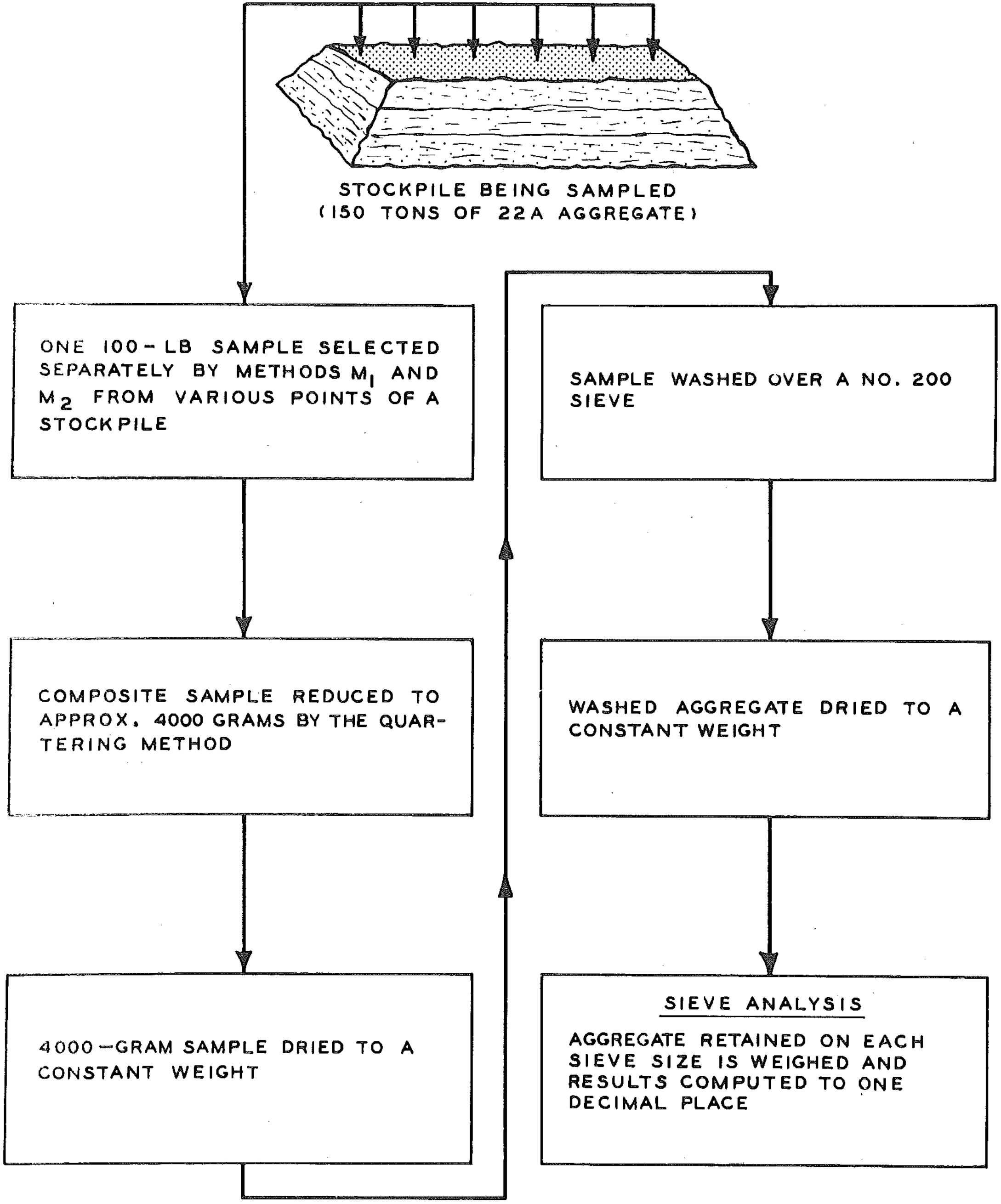


Figure 15. Flow chart for sampling and screening coarse aggregates.

4) Whether significant interactions (i.e., combined effects produced by several factors, exceeding the total of their individual effects) are occurring in the experimental work.

Experimental Procedure

A portable plant was selected, producing 22A aggregate for base and surface courses near Maple Rapids, Michigan. For this type of material the 1963 specifications required the following grading limits:

Sieve size	Percent passing (cumulative)
1-in.	100
3/4-in.	90 to 100
3/8-in.	65 to 85
No. 10	30 to 45
No. 200(1)	3 to 10

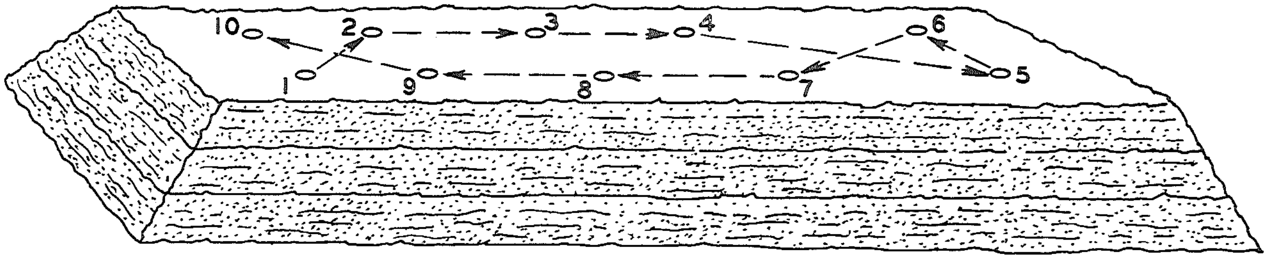
(1) Loss by washing.

The experimental design required three aggregate inspectors selected from a number of well-trained personnel from different locations, and three screening kits also selected from sieves available from the Testing Laboratory Division.

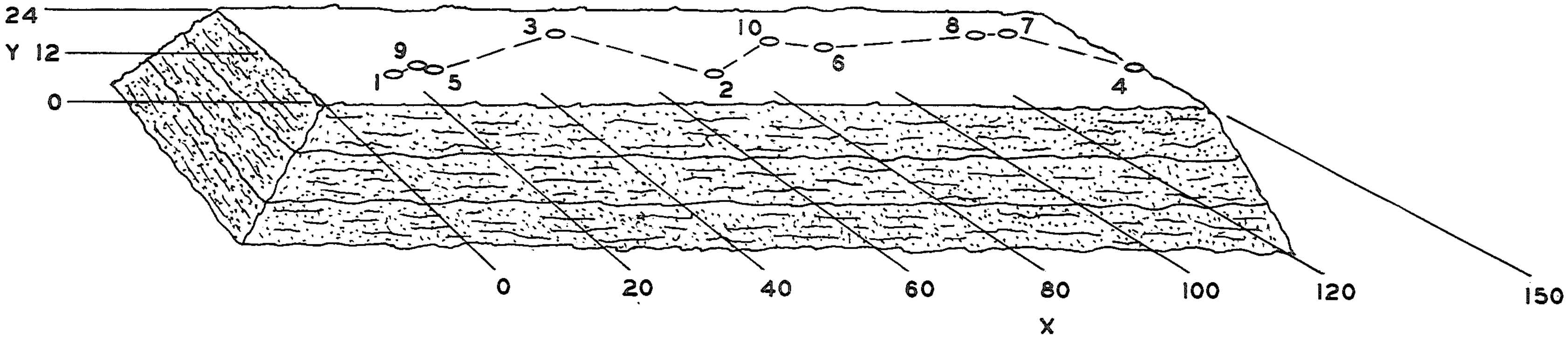
Each of the three aggregate inspectors used each screening kit equally often, and sampled and tested 60 samples for about four weeks of aggregate production. During the course of the experiment, a regular aggregate inspector informed testing personnel as to the status of the material being produced. The specifications require as a standard one complete gradation analysis for each 150 tons of coarse aggregate. Four or five tests per day cover the production from an average gravel plant. The regular inspector takes a representative sample by sampling from different areas of the stockpile and combining into a composite sample of about 100 lb of aggregate material. The composite or average sample is then reduced by quartering to a size suitable for testing for loss by washing (or passing No. 200) and sieve analysis. These field operations are illustrated by the flow chart in Figure 15.

In addition, the three inspectors performing the experimental testing obtained their samples from the same stockpile tested by the regular inspector. Both groups of inspectors took their samples prior to any additional

ONE STANDARD SAMPLING PROCEDURE CURRENTLY IN USE



RANDOM SAMPLING PROCEDURE



RANDOM SAMPLING RECORD

PROJECT NO. 63 G-123

SAMPLE NO. 21

DATE 7-8-64

TIME 11:00 AM

MATERIAL Base & Surface Agg.

SPECIFICATION 22A

PRODUCER A4

INSPECTOR I_2

KIT S_1

DESTINATION Stockpile (150' x 24')

LOCATIONS, FT		
SPOT	X	Y
1	18	7
2	74	8
3	55	18
4	150	11
5	26	9
6	100	14
7	139	19
8	131	19
9	23	10
10	91	16

Figure 16. Comparison of standard and random sampling procedures.

production being placed on the stockpile. The inspectors performing the experimental testing obtained their samples by the standard sampling procedure (shown schematically in Figure 16) and also by selected random sampling so that each sampling location from the stockpile would have an equal likelihood of being included in the composite samples. A random sampling pattern involving ten locations is also shown in Figure 16. Values for length X and width Y of the stockpile were computed from a table listing random numbers in decimal fractions and on actual stockpile dimensions of 150 by 24 ft. In both cases, the dotted lines show paths followed by the inspectors over the sampling areas.

Thus, two different sampling methods were included as part of the experimental design. To avoid unnecessary delay during the field testing, necessary space, equipment, and materials were furnished in a mobile laboratory truck parked near the project site. The experiment was conducted on accepted 22A aggregate without interference with the regular inspector's duties nor with aggregate plant operation. After following the sampling and testing procedures outlined in the design, the selected inspectors recorded their field observations on special form sheets. At the bottom of this form were coding letters for the inspector's name, testing equipment used, location of the material being sampled, the time of sampling, and the date. Then the data were punched on cards for computer processing.

Selection of Samples

Cost, labor, facilities, and time available all limited the sample size to 10 gradation tests for each combination of inspector-screening kit-sampling-method, as shown in Figure 17.

where:

n = 10 gradation analyses per cell
I_1, I_2, I_3 = three aggregate inspectors chosen at random from a large group
S_1, S_2, S_3 = three screening kits, also chosen from several that were available
M_1 = regular or standard method of sampling
M_2 = random method of sampling.

Survey records covering 22A aggregate indicate that for a randomized survey sample size of 10 gradation tests, a 20-percent risk of failing to detect genuine changes in the average value passing the 3/8-in. sieve was tolerated; with a 5-percent chance of erroneously recording nonexistent changes. On the other hand, for the same sample size, the risks of missing genuine changes in the average values passing Nos. 10 and 200 were close to 5 and 12 percent, respectively, with a probable error of 5 percent in recording nonexistent changes.

The experimental results are summarized below separately for each of the three sieve sizes involved. In all cases, extensive use was made of analysis of variance, a statistical technique for estimating how much of the total variation in gradation results can be significantly attributed to aggregate inspectors, screening sieves, sampling methods, or their interactions. The remaining variation due to other relatively non-significant causes is classed as residual variation, made up of inherent material and experimental variation.

	M_1			M_2		
	I_1	I_2	I_3	I_1	I_2	I_3
S_1	n	n	n	n	n	n
S_2	n	n	n	n	n	n
S_3	n	n	n	n	n	n

Figure 17. Experimental design involving three main factors: sampling methods, aggregate inspectors, and screening sieves.

Aggregate Passing 3/8-in. Sieve

The results of the analysis of variance (Table 6) indicate that the main effects of both M and I are highly significant— at the 1-percent level— but that there is no significant effect due to S. The results also show that the interaction MIS is significant at the 5-percent level. The possibility that this interaction might really occur is questionable since the 3/8-in. sieves seemed identical in physical appearance before and after cleaning in the field. On the other hand, the results may be interpreted better in terms of main effects M and I as summarized in Table 7 and shown graphically in Figure 18. All three inspectors obtained higher values when using the regular sampling method, but inspector I_3 was fairly stable with both methods, since he showed the smallest deviation from the grand average of 70.13 percent passing. This is reasonable because I_3 was the most experienced inspector of the three involved in the experiment.

TABLE 6
TABLE OF ANALYSIS OF VARIANCE FOR PASSING 3/8 IN. SIEVE

Nature of Effect	Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F	F Tests	
						F 0.05	F 0.01
Main Factors	M	162.64	1	162.64	15.00**	3.90	6.81
	I	106.64	2	53.32	4.92**	3.06	4.75
	S	21.20	2	10.60	0.98	3.06	4.75
Interactions Among Factors	MI	60.23	2	30.11	2.78	3.06	4.75
	MS	28.14	2	14.07	1.30	3.06	4.75
	IS	76.96	4	19.24	1.77	2.43	3.45
	MIS	124.96	4	31.24	2.88*	2.43	3.45
Replication	Residual	1756.14	162	10.84			
	Total	2336.92	179	13.06			

Legend:		
	**	Significant at the 1 and 5-percent levels (highly significant)
	*	Significant at the 5-percent level
	M	Sampling Methods
	I	Aggregate Inspectors
	S	Screening Kits

NOTE: Deviations are significant when they are large enough to make the assumption of equal performance unlikely. They are significant when the number in the "F" column is larger than the corresponding number in the "F 0.05" subcolumn, and highly significant when larger than the corresponding number in the "F 0.01" subcolumn.

Thus, inspectors and methods of sampling independently affected the gradation results passing the 3/8-in. sieve. A possible explanation for obtaining a lower average with a random sampling method is that a more representative sample can be expected by using this method as compared to regular sampling. Statistical theory states that a randomized sampling procedure increases the likelihood of getting samples representative of the lot, thus reducing or eliminating sampling as a source of bias. The investigation indicates that the bias is more likely to be on the higher side of the average when the standard sampling procedure is used. On the other hand, since both main factors M and I proved to be highly significant, interest would then center on how much of the overall variation in gradation results is due to discrepancy between the sampling methods and how much is due to discrepancy among different inspectors. An analysis of components of variance

TABLE 7
SUMMARY OF GRADATION TESTS ON 22A AGGREGATE PASSING 3/8-IN. SIEVE

Inspector	Item	Regular Method of Sampling, M_1	Random Method of Sampling M_2
I_1	No of Tests	30	30
	Avg. Gradation, %	70.52	67.57
	Std. Deviation, %	3.42	3.52
I_2	No. of Tests	30	30
	Avg. Gradation, %	71.79	69.33
	Std. Deviation, %	2.82	3.41
I_3	No. of Tests	30	30
	Avg. Gradation, %	70.92	70.63
	Std. Deviation, %	3.09	4.00
Total and Average	No. of Tests	90	90
	Avg. Gradation, %	71.08	69.17
	Std. Deviation, %	3.13	3.83

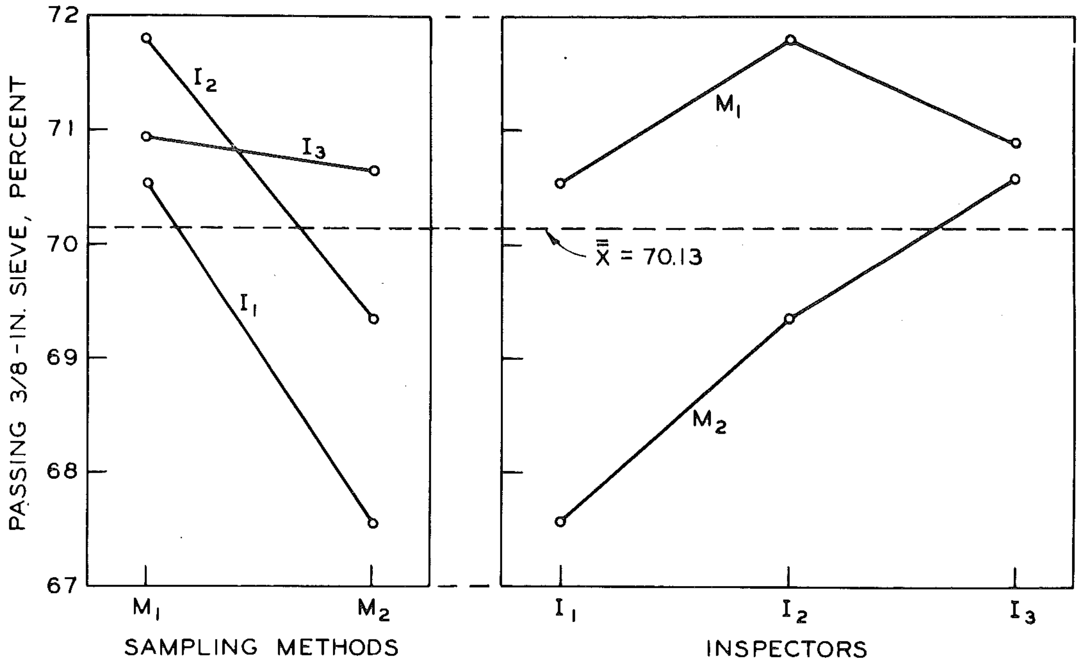


Figure 18. Significant variations due to sampling methods as well as inspectors.

gave an estimated value of 6 percent of the total variance attributable to sampling methods, and about 4 percent attributable to different inspectors. The remaining 90 percent of the total variance was due to residual variations including inherent material variation and experimental variations. Assuming that discrepancies of sampling methods and by inspectors could be eliminated or reduced by teaching better sampling procedures or developing more rigorous training programs, these estimated values indicate that much could be gained by such action.

TABLE 8

TABLE OF ANALYSIS OF VARIANCE FOR PASSING NO. 10 SIEVE

Nature of Effect	Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F	F Tests	
						F 0.05	F 0.01
Main Factors	M	97.98	1	97.98	10.67**	3.90	6.81
	I	39.10	2	19.55	2.12	3.06	4.75
	S	14.06	2	7.03	0.77	3.06	4.75
Interactions Among Factors	MI	25.29	2	12.64	1.38	3.06	4.75
	MS	3.61	2	1.80	0.20	3.06	4.75
	IS	280.20	4	70.05	7.63**	2.43	3.45
	MIS	98.81	4	24.70	2.69*	2.43	3.45
Replication	Residual	1487.32	162	9.18			
	Total	2046.36	179	11.43			

Legend:	
**	Significant at the 1 and 5-percent levels (highly significant)
*	Significant at the 5-percent level
M	Sampling Methods
I	Aggregate Inspectors
S	Screening Kits

Aggregate Passing No. 10 Sieve

The results of the analysis of variance (Table 8) indicate that the presence of interaction effects is significant enough to reduce the accuracy of the main effect comparisons. Here, the analysis must be supplemented by a detailed examination of the nature of such interactions. The combined influence (interaction) of inspectors with screening kits, significantly affecting the gradation results, is shown in Table 9 and Figure 19. Some combinations of the two factors (inspectors and screening kits) were better or worse than would be expected under existing conditions. Thus, from Figure 19, it appears that each inspector found a different kit that gave results with smallest deviation from the grand average of 43.20 percent passing. In fact, for better results, screening kit S_1 could have been assigned to inspector I_3, kit S_2 to inspector I_3, and kit S_3 to inspector I_1. However, under existing conditions, little could be gained by taking the trouble of assigning particular screening kits to particular inspectors.

TABLE 9
SUMMARY OF GRADATION TESTS ON 22A AGGREGATE PASSING NO. 10 SIEVE

Inspector	Items	Screening Kits		
		S_1	S_2	S_3
I_1	No. of Tests	20	20	20
	Avg. Gradation, %	41.27	43.88	42.55
	Std. Deviation, %	3.72	3.21	3.03
I_2	No. of Tests	20	20	20
	Avg. Gradation, %	43.96	41.91	45.09
	Std. Deviation, %	2.52	4.12	3.08
I_3	No. of Tests	20	20	20
	Avg. Gradation, %	45.40	42.79	42.07
	Std. Deviation, %	2.01	3.10	3.22
Total and Average	No. of Tests	60	60	60
	Avg. Gradation, %	43.54	42.86	43.23
	Std. Deviation, %	3.28	3.54	3.34

Possible reasons for such interactions might include: a) the inspector's practice of rearranging aggregate particles retained on the No. 10 sieve so they may pass; b) inspector's fatigue caused by shaking aggregate test samples by hand under field operating conditions; c) serious discrepancies in sieve openings over the same screening area caused by improper care of testing kits as used in the field.

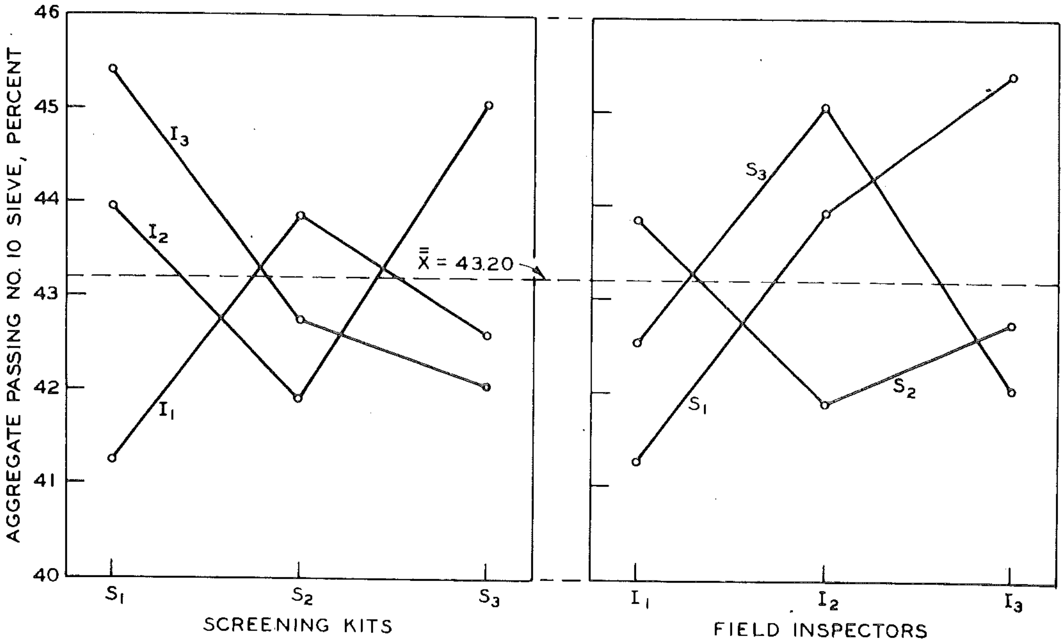


Figure 19. Combined influence (or interaction) of two factors (screening kits-inspectors) affecting gradation results.

On the other hand, the difference between the two sampling methods M_1 and M_2 was of sufficient magnitude in some cases to be of practical importance. As shown in Table 10 and Figure 20, relative performance of aggregate inspectors was not consistent for all kits, particularly when the two sampling methods are compared. For example, the best combinations were I_1-M_1-S_1, I_2-M_2-S_1, I_1-M_2-S_2, and I_2-M_2-S_3 because these were closest to the grand average of 43.20 percent.

When interactions are significant, the standard procedure for analyzing the components of variance is carried out for each factor separately. The results are as follows:

Source of Variance	Inspectors		
	I_1	I_2	I_3
Sampling Methods, percent	8	6	--
Screening Kits, percent	7	11	18
Residual Deviations, percent	85	83	82
Total	100	100	100

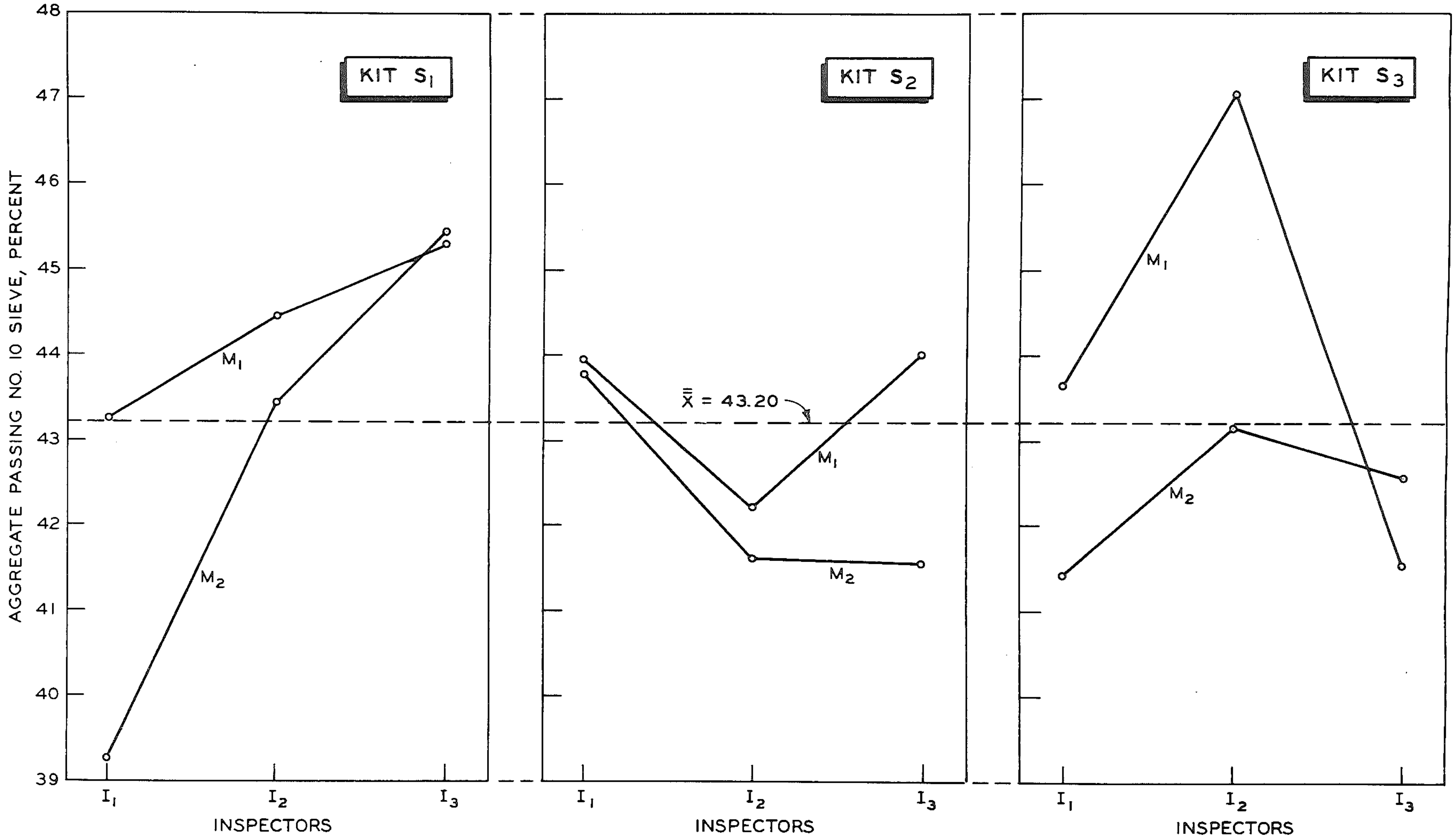


Figure 20. Relative performance of aggregate inspectors for individual kits and different sampling methods.

TABLE 10
SUMMARY OF GRADATION TESTS ON 22A AGGREGATE PASSING NO. 10 SIEVE USING TWO SAMPLING METHODS

Sampling Methods	Inspector	Item	Screening Kits		
			S_1	S_2	S_3
Regular Sampling Method M_1	I_1	No. of Tests	10	10	10
		Avg. Gradation, %	43.27	43.97	43.68
		Std. Deviation, %	3.50	3.43	2.97
	I_2	No. of Tests	10	10	10
		Avg. Gradation, %	44.47	42.20	47.05
		Std. Deviation, %	2.46	3.62	1.78
	I_3	No. of Tests	10	10	10
		Avg. Gradation, %	45.31	44.02	41.56
		Std. Deviation, %	2.12	1.81	3.56
Random Sampling Method M_2	I_1	No. of Tests	10	10	10
		Avg. Gradation, %	39.27	43.78	41.42
		Std. Deviation, %	2.84	3.16	2.79
	I_2	No. of Tests	10	10	10
		Avg. Gradation, %	43.44	41.62	43.12
		Std. Deviation, %	2.61	4.74	2.88
	I_3	No. of Tests	10	10	10
		Avg. Gradation, %	45.48	41.55	42.57
		Std. Deviation, %	2.00	3.68	2.93

For example, inspector I_1, always using the same calibrated kit and always sampling and testing in the same manner, might get gradation results with a constant bias, but the total variance of individual readings would be reduced by 15 percent. On the other hand, inspector I_3, being the most experienced, by always using the same calibrated kit might reduce the total variance by 18 percent. These estimated values would indicate how much might be gained by better calibration of screening kits or remedial training of aggregate inspectors selected for this experiment.

TABLE 11
TABLE OF ANALYSIS OF VARIANCE FOR PASSING NO. 200 SIEVE

Nature of Effect	Source of Variance	Sum of Squares	Degrees of Freedom	Variance Estimate	F	F Tests	
						F 0.05	F 0.01
Main Factors	M	2.45	1	2.45	2.21	3.90	6.81
	I	6.77	2	3.39	3.05	3.06	4.75
	S	0.26	2	0.13	0.12	3.06	4.75
Interactions Among Factors	MI	0.02	2	0.01	0.01	3.06	4.75
	MS	5.57	2	2.79	2.51	3.06	4.75
	IS	62.35	4	15.59	14.04**	2.43	3.45
	MIS	2.95	4	0.74	0.67	2.43	3.45
Replication	Residual	179.82	162	1.11			
	Total	260.19	179	1.45			

Legend: ** Significant at the 1 and 5-percent levels (highly significant)
M Sampling Methods
I Aggregate Inspectors
S Screening Kits

TABLE 12
SUMMARY OF GRADATION TESTS
ON 22A AGGREGATE PASSING NO. 200 SIEVE

Inspector	Items	Screening Kits		
		S_1	S_2	S_3
I_1	No. of Tests	20	20	20
	Avg. Gradation, %	7.58	9.07	7.93
	Std. Deviation, %	1.08	0.91	1.33
I_2	No. of Tests	20	20	20
	Avg. Gradation, %	8.28	8.34	9.38
	Std. Deviation, %	1.06	1.38	0.73
I_3	No. of Tests	20	20	20
	Avg. Gradation, %	9.27	7.99	7.92
	Std. Deviation, %	0.72	1.11	0.98
Total and Average	No. of Tests	60	60	60
	Avg. Gradation, %	8.37	8.46	8.41
	Std. Deviation, %	1.18	1.22	1.24

Aggregate Passing No. 200 Sieve

The analysis of variance (Table 11) indicates that highly significant interactions occurred between inspectors and kits during the experimental work. In fact, the combination of these two factors markedly affected the uniformity of gradation results as shown in Table 12 and Figure 21. The interpretation of these interrelated factors is somewhat similar to the results for material passing the No. 10 sieve. For example, it appears that some particular combinations (such as S_2-I_2, S_1-I_2, and S_2-I_3 or S_3-I_3) were better than others under experimental conditions. These interactions might occur for two principal reasons:

1) The difficulty in obtaining uniform effective sieve openings after the No. 200 mesh has been normally used in field work.

2) Variation among individual inspectors in their procedure of agitating and washing an aggregate test sample over a No. 200 sieve.

The nature of these interactions may be examined in detail in Table 13 and Figure 22 for each separate kit. Here, the best combinations were I_2-M_2-S_2, I_3-M_1-S_2, and I_2-M_1-S_1, because they were closest to the grand average of 8.40 percent passing the No. 200 sieve. In addition, these situations reveal that variations in sampling procedures were not significant, and thus such variations could be expected under typical field conditions.

Since interactions were significant in this case, the procedure for analyzing the components of variance are exactly the same as for aggregate passing the No. 10 sieve. The results are as follows:

Source of Variance	Inspectors		
	I_1	I_2	I_3
Sampling Methods, percent	--	--	--
Screening Kits, percent	23	15	28
Experimental Deviations, percent	77	85	72
Total	100	100	100

These estimated values would indicate how much more precision in gradation results might be attained by training inspectors to work alike, or by more careful calibration and maintenance of testing sieves.

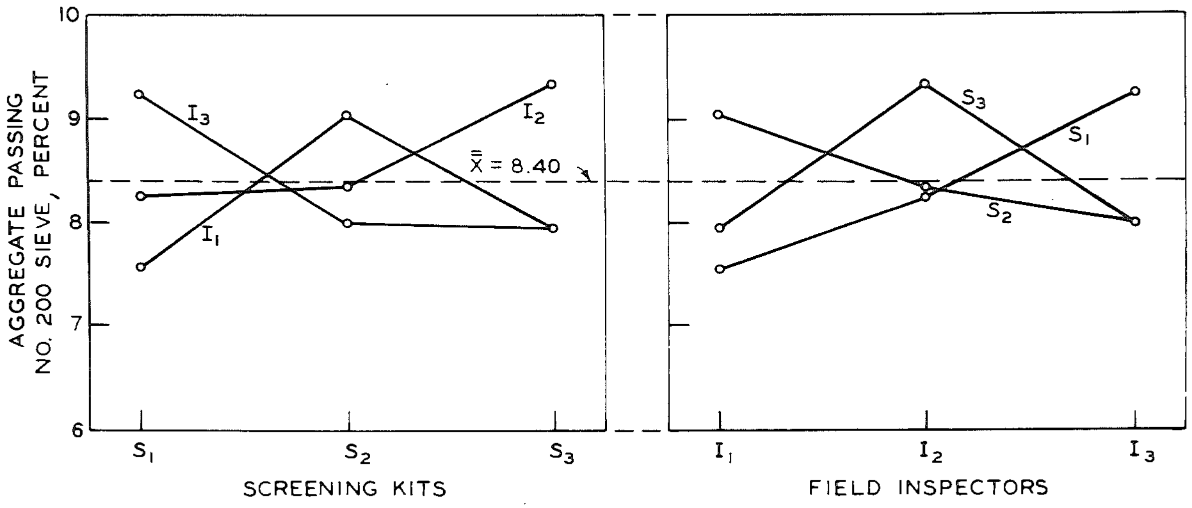


Figure 21. Combination of two factors (screening kits - inspectors) affecting uniformity of results.

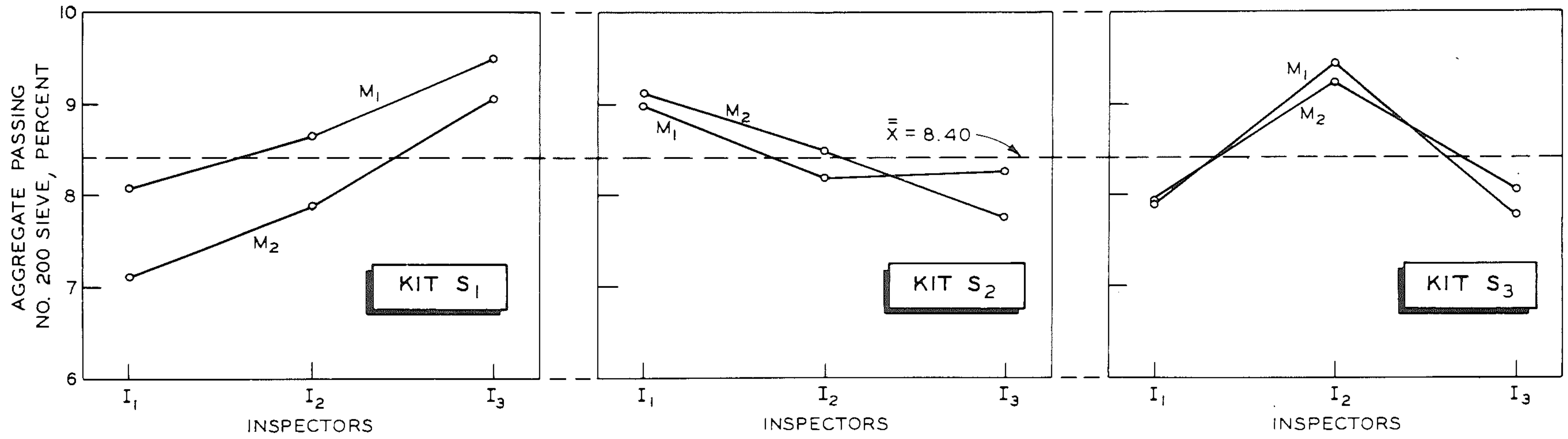


Figure 22. Interactions in various situations.

TABLE 13
SUMMARY OF GRADATION TESTS ON 22A AGGREGATE PASSING NO. 200 SIEVE, USING TWO SAMPLING METHODS

Sampling Methods	Inspector	Item	Screening Kits		
			S_1	S_2	S_3
Regular Sampling Method M_1	I_1	No. of Tests	10	10	10
		Avg. Gradation, %	8.06	8.99	7.90
		Std. Deviation, %	1.06	0.76	1.30
	I_2	No. of Tests	10	10	10
		Avg. Gradation, %	8.66	8.18	9.46
		Std. Deviation, %	1.28	0.88	0.62
	I_3	No. of Tests	10	10	10
		Avg. Gradation, %	9.49	8.24	7.79
		Std. Deviation, %	0.76	0.93	0.85
Random Sampling Method M_2	I_1	No. of Tests	10	10	10
		Avg. Gradation, %	7.09	9.14	7.95
		Std. Deviation, %	0.91	1.07	1.44
	I_2	No. of Tests	10	10	10
		Avg. Gradation, %	7.89	8.49	9.29
		Std. Deviation, %	0.63	1.79	0.85
	I_3	No. of Tests	10	10	10
		Avg. Gradation, %	9.04	7.74	8.04
		Std. Deviation, %	0.64	1.26	1.13

Control Charts for 22A Aggregate

For detecting process variation during continuous production of 22A aggregate, control charts for a regular inspector and three inspectors who worked under experimental conditions are shown in Figure 23. The regular inspector and the inspectors performing the experiment obtained their samples at the same time from the same stockpile. Since the No. 10 sieve occasionally causes trouble to producers, this sieve was selected to illustrate relative performance of the inspectors in relation to acceptability of the material. The control charts show daily average results passing the No. 10

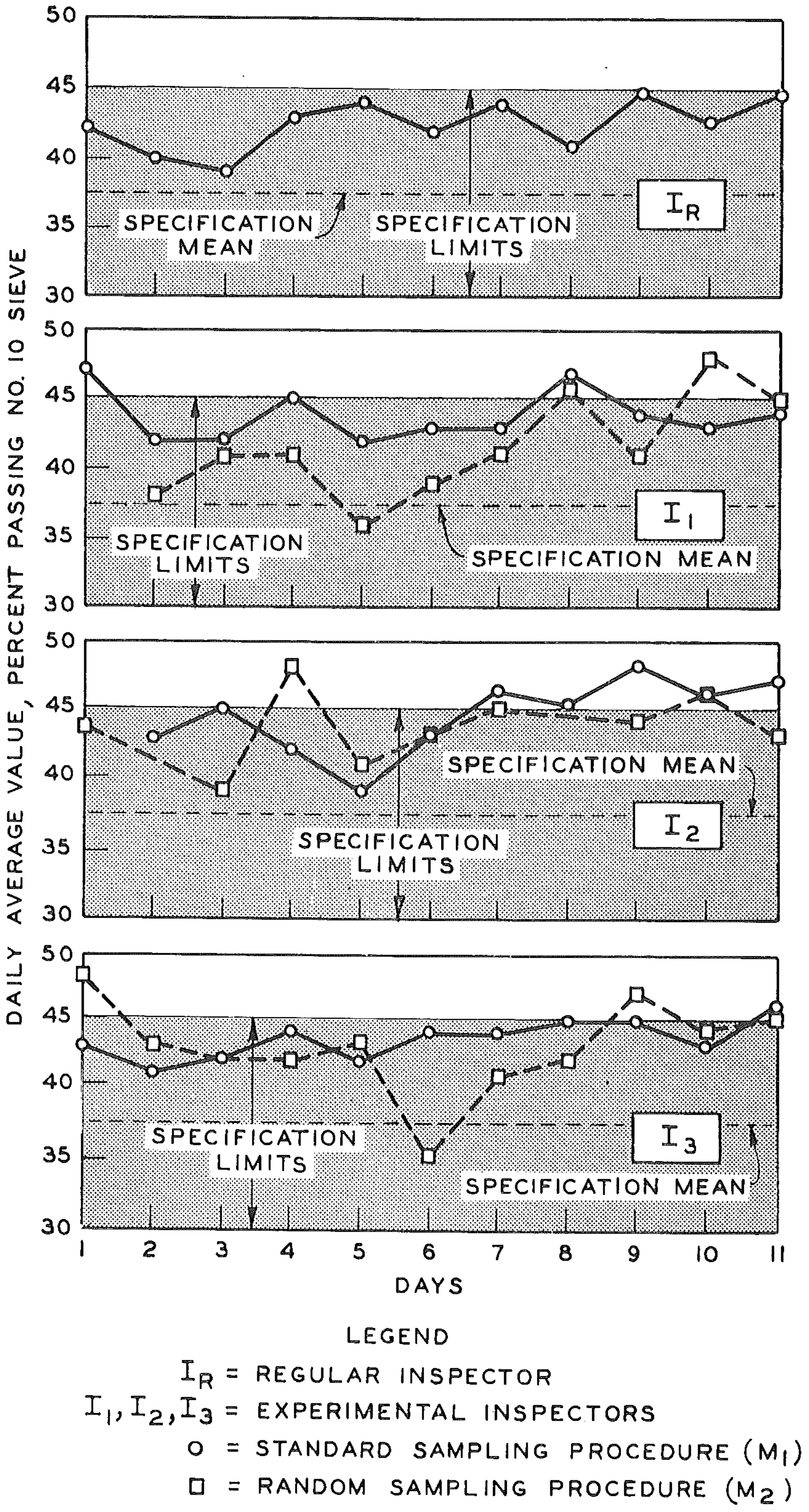


Figure 23. Relative performance of aggregate inspectors using standard and random sampling methods.

sieve during 11 consecutive days of acceptable aggregate production. These charts disclose the following:

1) Daily average values fell consistently above the specification mean. In addition, all four charts show an increasing trend, suggesting that production conditions were changing and that an investigation of production problems might well be justified at this point.

2) Because of this increasing trend, inspectors I_1, I_2, and I_3 probably would have rejected about 20, 40, and 10 percent, respectively, of the total aggregate material already accepted by the regular inspector I_R. Here, definite decision rules might be desirable regarding the acceptability of material whenever successive average values show increasing or decreasing trends on the same side of the specified mean.

3) In general, lower average values were obtained with the random sampling method than with the standard procedure. Furthermore, the difference between the two sampling methods was remarkably consistent for inspector I_1. This finding supports the results of the analysis of variance, as previously explained.

Results of the Field Experiment

Briefly, the results of the field study are as follows:

1) Inspectors and methods of sampling independently affected gradation results passing the 3/8-in. sieve. An analysis of components of variance gave an estimated value of 4 percent of the total variance attributable to different inspectors, 6 percent attributable to sampling methods, and the remaining 90 percent attributable to inherent material and experimental deviations.

2) The test data on passing No. 10 sieve showed the presence of significant interaction effects among the main factors involved in the experiment. Analysis of components of variance was carried out for each inspector separately with the vollowing results:

a) Variance due to different ways of selecting the sample ranged from 0 to 8 percent.

b) Variance due to different screening sieves ranged from 7 to 18 percent.

c) The remaining variance of 85 percent is attributable to inherent material and experimental deviations.

3) Similarly, the test data on passing the No. 200 sieve were affected by significant interactions between inspectors and kits during the experimental work. Analysis of components of variance was carried out for each inspector separately with the following results:

a) Variance due to different screening kits covered 15 to 28 percent.

b) Variance due to inherent material and experimental deviations was 72 to 85 percent.

c) Variations in sampling procedures were not significant.

4) Control charts for 22A aggregate disclosed at least four significant features not covered by current conventional records-keeping methods, but nevertheless inportant in an efficient sampling acceptance procedure:

a) Daily average values fell consistently above the specified mean.

b) Daily average results showed a gradual upward trend.

c) Test inspectors would have reached different decisions on acceptance or rejection of the material already accepted by the regular inspector.

d) In general, higher values were obtained with the standard method of sampling than with the random sampling procedure. Also, inspector I_1 appeared to maintain a consistent difference between the two sampling methods.