PB91-176321

U.S. Department
of Transportation

**Federal Highway
Administration**

Publication No. FHWA-TS-89-038
October 1990

Pavement Construction

Quality Assurance in Highway Construction

Office of Implementation
Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, Virginia 22101-2296

Implementation
advancing highway technology

FOREWORD

This report contains the reprints of six articles on the subject of quality assurance that have appeared in past issues of "Public Roads" magazine.

The quality of the highway product has always been a major concern to highway engineers and contractors. Quality assurance in its broad application relates to this concern of obtaining the quality of construction necessary for successful performance. It encompasses design, production, sampling, testing, and decision criteria.

These articles will be of interest to engineers and technicians involved in pavement design, construction, and rehabilitation.

Sufficient copies of this report are being distributed by Federal Highway Administration (FHWA) memorandum to provide one copy to each FHWA Region and Division offices and one copy to each State highway agency. Direct distribution is being made to the Division offices. Additional copies for the public are available from the National Technical Information Service (NTIS), 5280 Port Royal Road, Springfield, Virginia 22161.

Director, Office of Implementation

NOTICE

Technical Report Documentation Page

1. Report No.
FHWA-TS-89-038

2
PB91-176321

3. Recipient's Catalog No.

4. Title and Subtitle

Quality Assurance in Highway Construction

5. Report Date
October 1990

6. Performing Organization Code

7. Author(s)
McMahon, Halstead, Baker, Granley & Kelly

8. Performing Organization Report No.

9. Performing Organization Name and Address
Federal Highway Administration
"Public Roads" Magazine
6300 Georgetown Pike
McLean, VA 22101

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

12. Sponsoring Agency Name and Address

Office of Implementation
Federal Highway Administration
6300 Georgetown Pike
McLean, Virginia 22101 - 2296

13. Type of Report and Period Covered

Final Report

14. Sponsoring Agency Code

15. Supplementary Notes

Reprint from "Public Roads" Magazine, Volume 35, Numbers 6 - 11

16. Abstract

This report contains the reprint of six articles on the subject of quality assurance that have appeared in the past issues of "Public Roads" magazine.

The articles are divided into the following:

1. Introduction and Concepts
2. Quality Assurance of Embankments and Base Courses
3. Quality Assurance of Portland Cement Concrete
4. Variations of Bituminous Construction
5. Summary of Research for Quality Assurance of Aggregate
6. Control Charts

17. Key Words

Quality Assurance, Pavement
Construction, Pavement Materials,
Asphalt Concrete Pavement,
Portland Cement Concrete Pavements.

18. Distribution Statement

No restrictions. This document is available
to the public through the
National Technical Information Service,
Springfield, Virginia 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	70	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

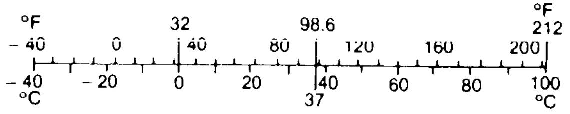
Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimetres	mm
ft	feet	0.305	metres	m
yd	yards	0.914	metres	m
mi	miles	1.61	kilometres	km
		AREA		
in^2	square inches	645.2	millimetres squared	mm^2
ft^2	square feet	0.093	metres squared	m^2
yd^2	square yards	0.836	metres squared	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	kilometres squared	km^2
		VOLUME		
fl oz	fluid ounces	29.57	millilitres	mL
gal	gallons	3.785	litres	L
ft^3	cubic feet	0.028	metres cubed	m^3
yd^3	cubic yards	0.765	metres cubed	m^3

NOTE: Volumes greater than 1000 L shall be shown in m^3.

Symbol	When You Know	Multiply By	To Find	Symbol
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5(F-32)/9	Celcius temperature	°C

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimetres	0.039	inches	in
m	metres	3.28	feet	ft
m	metres	1.09	yards	yd
km	kilometres	0.621	miles	mi
		AREA		
mm^2	millimetres squared	0.0016	square inches	in^2
m^2	metres squared	10.764	square feet	ft^2
ha	hectares	2.47	acres	ac
km^2	kilometres squared	0.386	square miles	mi^2
		VOLUME		
mL	millilitres	0.034	fluid ounces	fl oz
L	litres	0.264	gallons	gal
m^3	metres cubed	35.315	cubic feet	ft^3
m^3	metres cubed	1.308	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.205	pounds	lb
Mg	megagrams	1.102	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celcius temperature	1.8C + 32	Fahrenheit temperature	°F



* SI is the symbol for the International System of Measurement

(Revised April 1989)

Table of Contents

List of Figures

Part 1 - Introduction and Concepts

Part 2 - Quality Assurance of Embankments and Base Courses

List of Figures (continued)

Part 3 - Quality Assurance of Portland Cement Concrete

Part 4 - Variations of Bituminous Construction

List of Figures (continued)

Part 5 - Summary of Research for Quality Assurance of Aggregate

Part 6 - Control Charts

List of Tables

Part 2 - Quality Assurance of Embankments and Base Courses

Part 3 - Quality Assurance of Portland Cement Concrete

List of Tables (continued)

Part 4 - Variations of Bituminous Construction

Part 5 - Summary of Research for Quality Assurance of Aggregate

List of Tables (continued)

Part 6 - Control Charts

Reprinted from
PUBLIC ROADS
A Journal of Highway Research
Vol. 35, Nos. 6–11

Quality Assurance in Highway Construction

Part 1—
Introduction and Concepts

**Reported by THURMUL F. McMAHON,
Principal Quality-Assurance Research Engineer,
and WOODROW J. HALSTEAD,
Chief, Materials Division**

Quality assurance in its broad application relates to the overall problem of obtaining the quality of construction necessary for a product to perform the functions intended. It encompasses design, production, sampling, testing, and decision criteria.

The quality of the highway product has always been a major concern to highway engineers and contractors. Traditionally, quality has been attained primarily through skills of individual engineers. When such skills are properly applied, satisfactory highway quality is obtained. However, as the speed of construction and the volume of materials to be handled increased, the traditional system became subject to breakdown. Breakdown occurs when the speed of testing does not keep pace with the speed of construction. Additionally, engineering duties have increased to the extent that engineers must spread their talents over broad areas, and many quality assurance activities have been delegated to those whose skills and experience are often inadequate for on-the-spot judgments. Moreover, legal requirements for documented evidence of specification compliance create problems.

As the Interstate program moved into its full construction phase, it became evident that the traditional quality assurance procedures were subject to criticism and that new concepts were needed. Accordingly, in 1963, the Public Roads Director of Research and Development appointed a task force to study the problem and develop a cooperative State-Public Roads research effort to improve quality assurance methods in highway construction.

The discussions and date presented here are an interpretative summary of the research progress in this area; some of the discussions already have been released by the Office of Research and Development *(1)*.[1] The reader should be aware that this article pertains to a Research and Development program—not to Public Roads policy. All the proposals presented will be carefully evaluated and only those proven to be workable under actual highway-construction conditions will be adopted as parts of State or Public Roads specifications and policy.

Basic Problems of Quality Assurance

Reduced to its simplest terms, quality assurance of highway construction requires proper answers to the following

---

[1] Italic numbers in parentheses identify the references listed on p. 8.

three questions: (1) What do we want? (2) How do we order it? (3) How do we determine that we got what we wanted?

Answers to the first question encompass the total body of research, development, engineering technology, and experience. All these combine to define needs with respect to materials, properties, and design characteristics of the highway component.

Answers to the second question depend on the manner in which the details are spelled out in specifications—specific characteristics that must be controlled, needs with respect to qualitative level, and uniformity of the product from item to item.

Answers to the third question depend on the precision and accuracy of test methods as well as on the time required to perform the tests. Testing time often controls the number of measurements that can be made available for use in decisionmaking. More importantly, the relation of the characteristic, or property, measured by the test to the service performance of the completed component is a major consideration, which often is known only empirically, if at all.

Traditional Quality Assurance

Many specifications used today in highway construction are, in fact, recipes rather than specifications. They spell out in detail the operations of the contractor, the equipment he must use, and the desired end product he must produce. These traditional specifications have come about because adequate quality definitions and test methods pertaining to quality of the end product are lacking. When specifications do attempt to define required quality, the specified values for characteristics are often those obtained through judgment and experience. Tolerances for such characteristics seldom reflect the true needs and capabilities of the construction process or of the available materials.

When traditional specifications are combined with the skills of engineers, the complete cooperation of contractors, and the desire of everyone to do a good job, there is no doubt that a good highway can be built. However, inspectors and engineers must be capable of recognizing good materials and construction, without relying solely on quality measurements. Under most of the present procedures, one periodic sample is taken. This sample—assumed to be representative of the material or construction—is tested, and the test result is recorded as the value of the measured property, or characteristic. If the test result is within the stated tolerances, the material *passes* and is accepted. If the test result is not within the stated tolerances, the material or construction fails to *pass*. Engineering judgment must then be applied and a decision made as to whether the material should be retested or whether it may be said to *substantially comply* because the specification deviation will cause little impairment of performance.

Even though a quality assurance system that is based on engineering judgment is workable under proper conditions, the practice is difficult to define in legal or contractual terms. *Substantial compliance* has not been quantitatively defined, and the degree of acceptable variation will differ from engineer to engineer and from job to job.

To further complicate the problem, sampling and testing errors are often so large that the true variations of the materials or construction may be obscured. Some tests may not truly measure quality of the finished highway.

Improvement in quality assurance of highway construction accordingly entails:

- Development of realistic quality criteria.
- Development of valid quality tests.
- Development of valid decisionmaking rules.
- Quantification of substantial compliance.

New Developments in Quality Assurance Procedures

Statistical concepts are the most promising tools for the solution of many quality-assurance problems in highway construction. Other industries have been using statistical concepts in process control and acceptance. In fact, much of the development in this field was pioneered by the Department of Defense in its procurement program during World War II. Because of the nature of the highway industry, some of the methods must be modified, but the concepts are basic to any industry.

The science of statistics is a versatile tool. In situations requiring decisions concerning contractual items that are based on samples, statistical concepts allow varied acceptable solutions. Rules for each decision must be carefully defined and followed, but different rules can be formulated for each of the many conditions encountered. Decisions can be made with an established degree of confidence. The degree of confidence required for each decision can be correlated with the criticalness of the decision to the quality of the end product, and the rules formulated accordingly.

Test methods are continually being developed for better and more rapid measurement of quality. The greatest advance in new methods of testing has been in the nuclear field. The nuclear moisture-density gage (figs. 1 and 2) has been proved to be a fast, accurate method of measuring the moisture and density of compacted materials. Nuclear methods (fig. 3) of measuring density and asphalt content of bituminous pavements are showing considerable promise. Seismic methods of measuring compaction are also being developed. Sonic equipment is being used to test welds, and sonic methods of measuring the moduli of concrete have been in use for several years, but have not been widely accepted. Electronic equipment, using the principles of resistivity and magnetism, has been developed to check the placement of steel in concrete and to measure the thickness of pavement components.

Rapid nondestructive tests such as those cited will provide better quality control and make quality measurements available in the future.

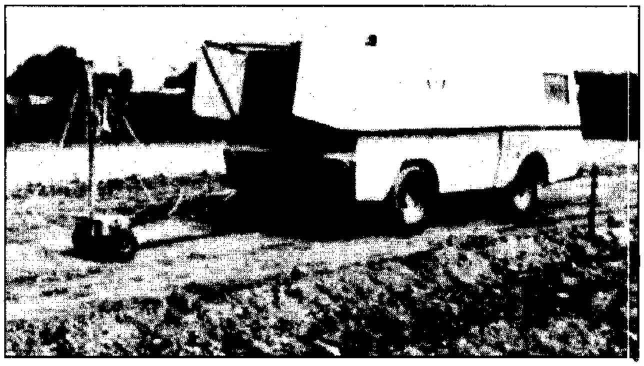


Figure 1.—The nuclear ROADLOGGER used for moisture-density determinations in compacted embankments.

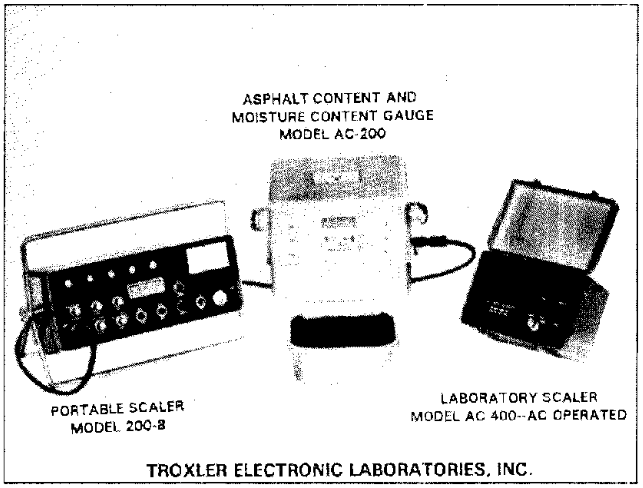


Figure 2.—Moisture-density determination in compacted embankment using portable nuclear gage.

Figure 3.—Portable nuclear gage in field test to determine density of bituminous base.

Figure 4.—STRINGLINE wire guidance system for controlling the placement of bituminous material.

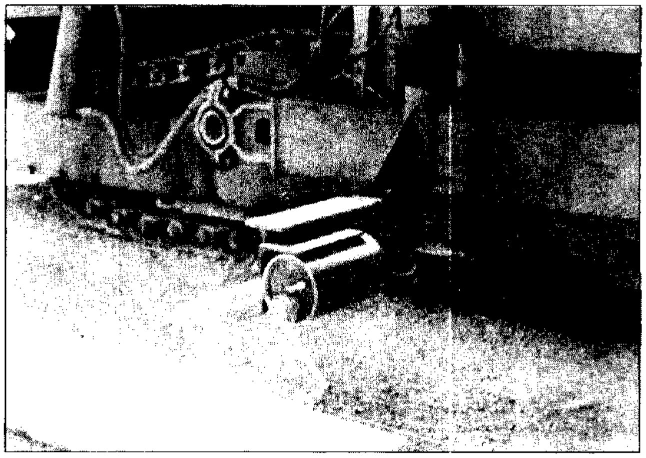


Figure 5.—SKI and wire guidance control used to provide smooth placement of pavement.

Through the work of its different committees, the American Society of Testing & Materials (ASTM) is advancing the state of the art of quality measurement by developing precision statements for standard tests. These statements will provide a basis for evaluating the work of inspectors and laboratory technicians and should decrease testing errors.

Other aids to better quality products are automated processing plants with direct output printout. These plants provide not only automatic control, but also adequate documentation to check output for pay quantities. However, automated control is no guarantee of a quality product. One must know what to control and how precise the control must be before the benefits of automation can be attained. One area in which automation is producing dramatic results is that of surface-variation control. The *Stringline* (fig. 4), a wire guidance system to control vertical variations in concrete placement, and other guidance methods (fig. 5) have greatly improved the riding quality of pavements.

Advantages of Statistical Concepts

One significant problem in quality assurance is that of communication. Definite instructions concerning the materials and construction desired, methods to be used for determining compliance, and conditions under which payment will be made, must be given to contractors. These instructions must be explicit so that contractors, engineers, lawyers, and auditors can interpret them in only one way. The Office of Research and Development, Bureau of Public Roads, recommends that statistical concepts be incorporated in the specifications for highway construction to improve communications.

The proper use of statistical concepts will provide the following requisites:

- Statement of concise quality requirements.
- Development of valid tolerances based on the capabilities of process, sampling, and testing methods.
- Delineation of responsibility for process control and acceptance.
- Development of valid sampling plans as a basis for decisionmaking.
- Establishment of precise decision critiera.
- Development of valid proportional-payment schedules.

Stating quality requirements.

In the writing of specifications, statistical concepts can be used to express quality requirements as target values for which contractors are to aim, and to specify compliance requirements as plus and minus tolerances. Tolerances from the target value, prescribed by design needs, can be based on statistical analyses of the variations in materials, processes, sampling, and testing existing in current construction practices. Such tolerances are realistic and enforceable. They take into account all the normal causes of variation and allow for the expected distribution of test results about the mean. Provisions can be made both for control to the stated level and for control of the variation from this level.

Research by the States is being undertaken to define realistic tolerances on quality requirements. From this research, it is known that test measurements on characteristics of highway materials or construction form a definite pattern grouping around a central value called the mean. The grouping indicates that test measurements in highway construction can be described in the same terms as test measurements in other industries. The measurements group around the central value in a symmetrical pattern, thereby allowing the use of statistics based on the familiar bell-shaped *normal curve*. Although some slight variation from the symmetrical curve may occur, especially when the number of test results is

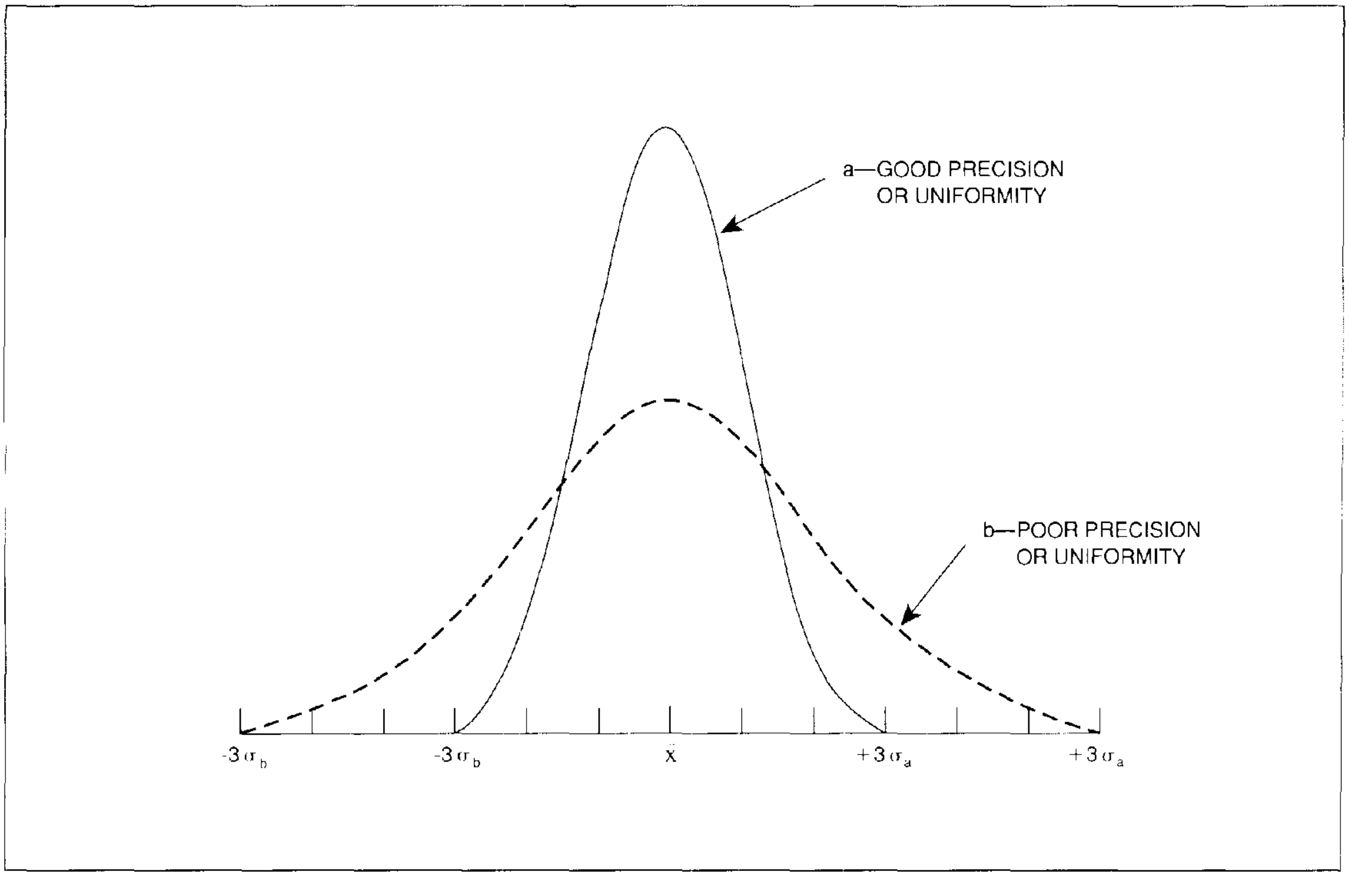


Figure 6.—Normal distribution curves.

small, the error in assuming normal distribution of population measurements usually will not be large. If the curve is decidedly asymmetrical, skewed to the right or left, then something other than normal distribution theory must be used in the analysis.

Even though curves are normal, they may not look alike. Those with a small standard deviation will be tall and narrow, whereas those with a large standard deviation will be short and broad. (See fig. 6.) The tall narrow curve indicates good product uniformity or measurement precision; the short broad curve indicates poor uniformity or precision.

The assumption of a normal distribution when warranted, permits the use of estimated relationships of mean and standard deviation to establish realistic specification tolerances for selected sample sizes. Such tolerances can be established by statistical analysis, together with engineering judgment, according to the degree of control needed for permissible construction risks and the economics of testing. The number of test results on which the compliance decision is based directly influences the latitude that must be given to the contractor. Often, because of the small number of tests that can be made economically, the tolerances must be wider than would seem desirable.

These relations may be stated as follows:

$$T_s = \frac{Z}{\sqrt{n}}$$

Where,

T_s is the tolerance to be allowed on each side of the target value.

Z is a standardized factor equal to $(X - \bar{X})/\sigma$ that relates to the area under the normal curve for the desired confidence of decision.

n is the number of tests to be made (sample size).

Statistical concepts for quality assurance of highway construction are based on the laws of probability; consequently, these laws must be allowed to function. One of the most important requirements for proper functioning is that the data be selected by random sampling. A true random sample is one in which all parts of the whole have an equal chance of being chosen for the sample. A table of random numbers is the best device for achieving a strictly random sample, but another method of chance, such as dice, the tossing of several coins, or a wheel of chance, often will suffice in highway work. The principal requirement is that the sample not be biased by a set selection pattern or by an inspector seeking either good, bad, or representative parts for sampling.

In addition to the laws of probability, another concept, *lots*, is essential to the proper application of statistics to quality control and acceptance sampling of highway construction. A *lot* is a uniquely identified, homogeneous portion of material or construction about which a decision is to be made. The size of the *lot* may vary depending on the economics of rejection and on sampling and testing costs. The lot size must not impose a severe hardship on the contractor who encounters a rejection—the smaller the *lot* the better the contractor's position. However, small *lots* entail more sampling and testing by the State—the larger the *lot* the better the State's position. Therefore, *lot* size must be a compromise equitable to both.

Production quality control

The application of statistical concepts to highway construction allows a definite assignment of responsibility for product quality. The contractor strictly is responsible for providing quality materials and construction; the State has the prerogative of acceptance sampling and testing.

Each contractor or supplier should have a statistical quality control program that will assure his meeting the acceptance requirements of the State. Such a quality control program can be patterned after the control currently exercised by the State or it can be considerably different. Much of the control of materials and construction can be accomplished by tests, usually called indicator tests, that are somewhat simplified. These tests are less precise but more rapid than the standard tests. When proper correlation has been established, a sufficient number of indicator-test results will provide control that is as good as fewer results from more precise tests.

Control charts are among the most useful tools in production quality control. These charts, on which test results are plotted, are simple line graphs of the required quality level and of the allowable variations from this level. They pictorially present data so that everyone concerned can see the results and readily observe trends that may affect quality.

Control charts depict data in several ways, and they can be of a simple design in which the target value is used as the axis and the specification limits as the control limits. Such charts show the variation of individual values or averages with respect to the actual specifications. However, when the mean, standard deviation, and the range of the material or process can be computed from a sample, average and range charts should be used.

The average, $\bar{X}$, chart shows variations in the averages of test results. A central line and upper and lower control-limit lines are used. The range, R, chart shows variations in the ranges of test results. It also has a central line and upper and lower control limits. Construction of these charts is described in any good quality control text.

If the average, $\bar{X}$, chart is being used to control current production, a sample of n items is taken from the process at random intervals and a quality measurement made on each item. The average of these measurements is then computed and plotted on the chart. As long as the sample averages neither fall outside the control limits nor show any nonrandom variation within the limits, the process is deemed to be in control with respect to its central tendency or *target value*.

When a range, R, chart is being used to control current output, the range of a sample of n items is computed and plotted on the R chart. If the sample ranges neither fall outside the control limits nor show any nonrandom variation within the limits, the process is considered to be in control with respect to its variability. The $\bar{X}$ and R charts must be used together to assure control of both level and variation of quality. Examples of Average, $\bar{X}$, and Range, R, charts are shown in figures 7 and 8.

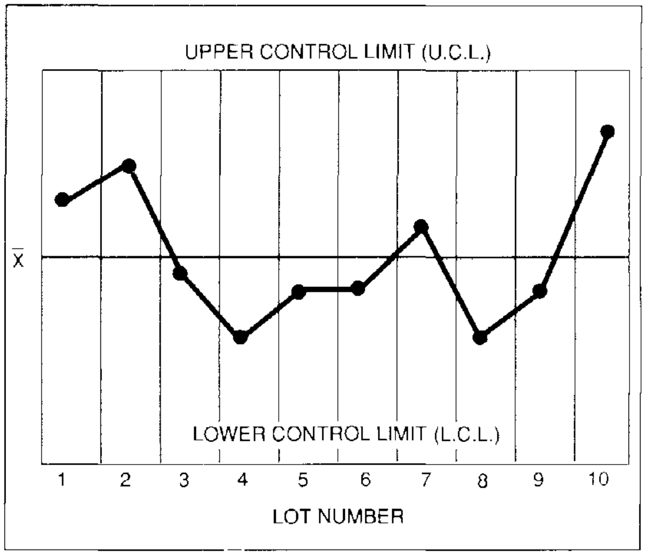


Figure 7.—Average, $\bar{X}$, control chart for n samples per lot.

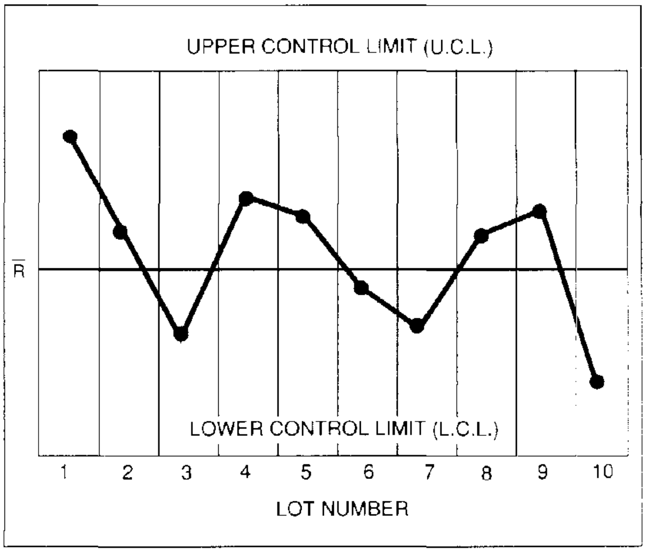


Figure 8.—Range, R, control chart for n samples per lot.

Acceptance procedures

For highway construction, the State may elect to use the results of supplier's or contractor's quality-control programs to accept material or construction. However, the usual procedure in buyer-seller relations is for the buyer to establish independent acceptance plans for each item of material or construction. An acceptance plan designates lot size; where and when to sample, on a random basis; numbers of samples to be taken; method of test to be used in the quality measurement on each specified characteristic of the sample; and, based on the test results, procedure for making a decision. An acceptance plan may be a simple statement or a complicated system in which many steps must be taken before a decision can be made. Examples of sampling plans will be included in subsequent installments of this article.

When decisions are based on a sample, a basic truth must be accepted: There is a certain risk that the decision is incorrect because the sample does not truly describe the total of the material. One advantage of the statistical approach is the ability to design a sampling plan in which the probability of acceptance of poor material, the β risk, and the rejection of good material, the α risk, are known. When *good* and *bad* materials have been defined and the risks to be taken agreed upon, the number of samples required to make a decision compatible with the risk probabilities can be calculated. These relations and the methods for establishing an operating characteristic curve, which denotes the probability of acceptance for intervening qualities of a product, can be found in any good quality control text.

Summary of Research Effort

During the past 4 years, the Office of Research and Development, Bureau of Public Roads, has actively promoted the following five-point program of research in quality assurance in highway construction:

- Awakening the highway industry's interest in the utility of the statistical approach to quality control and acceptance testing.
- Developing guides for research that would yield statistical data for writing acceptance specifications.
- Planning and coordinating a nationwide program of research in applying statistical methods to highway construction.
- Gathering and analyzing data and disseminating research findings.
- Designing and implementing projects by which the findings of the research program can be evaluated.

This effort is basically State research financed with Highway Planning and Research (H.P. & R.) funds. Many of the studies have been conducted according to guidelines established by Public Roads Task Force; others have followed plans developed by State personnel.

Early in the research program it was realized that little data were available for use in establishing quality levels and variations in highway construction on a statistically valid basis. Therefore, a concerted effort was initiated to measure quality and its variations in terms of existing criteria. Participating State highway departments have been measuring the level and variability of quality in their construction. To date, 28 States have conducted studies funded under H.P. & R. contracts and seven others have been investigating construction in State-funded studies.

The objective in the formulation of all studies was to produce compatible information that could be used throughout the Nation. A booklet of guidelines (2) was prepared and distributed to the States for use in planning their projects. A method of obtaining statistically valid data for an analysis of variance to isolate the components of variance was outlined in a suggested research plan. The plan permits overall variance to be divided into material or process variance, sampling variance, and testing variance.

According to the research data received from the States, 50 percent or more of the overall variance could be attributed to sampling and testing in some of the studies. Results showing this magnitude of sampling and testing error indicate that a concerted effort should be exerted by each highway department to train inspectors and laboratory technicians.

Also, according to the research data, which has been statistically analyzed to determine the percentage of present construction that complied with the levels and limits of current specifications, a considerable portion of the construction is shown to be outside the limits defined by the specifications. In fact, as much as 30 percent of some construction, considered to be completely acceptable under current control procedures, may be outside the stated limits. This variation from the specifications, in part, reflects the errors of sampling and testing, but there are indications that many of the present limits do not reflect valid allowances for the variable materials and processes used in highway construction.

Supplementing the State research effort, the Public Roads Office of Research and Development began a contract-research program in 1963 to further the development of statistical quality-control applications to highway construction. Many aspects of the task force's research plan were based on the results of the initial study in which the contractor evaluated the choice of concepts available and pointed to the priority areas for study. The study conclusions were presented in an unpublished report entitled *A Plan for Expediting the Use of Statistical Concepts in Highway Acceptance Specifications*. Two subsequent contracts provided valuable information concerning the level and variation of quality in base and subgrade construction.

A review of the Public Roads Standard Specifications for *Construction of Roads and Bridges on Federal Highway Projects* (FP–61) was conducted by another contractor. The final report on the contract was later used to develop a *futurized* revision of FP–61—the first attempt at writing complete specifications using statistical concepts wherever feasible.

The *Futurized Revisions of Federal Project Specifications* was never intended for use in highway construction, and distribution of the document has not been widespread. However, it has been reviewed by many outstanding highway engineers and by committees of the American Road Builders Association (ARBA) and other organizations. Most of the comments received have been favorable to the concepts incorporated in the specifications, but some disagree with methods of accomplishment and with items other than those that were treated statistically. The statistical applications embodied in the *Futurized Revision of Federal Project Specifications* have been proved to be sound and are the basis of many specifications now being written.

Subsequent information obtained from the States' research studies and Public Roads' in-house research has been used in the development of statistically based research specifications for construction of embankments, bases, and bituminous pavements. These specifications have been studied and discussed by many engineers associated with highway construction. It is evident from the comments received that some of the ideas presented are still not completely acceptable to the industry. Objection has been voiced to the complete delegation of quality control responsibility to the contractor and to the reduced payment schedules for nonconforming materials and construction. Primarily, the differences of opinion concern the degree of responsibility and the amount of reduced payment.

Undoubtedly, changes in present contractor-State relations are needed to fully implement the statistical approach to specifications. These changes must establish end-result requirements that can be measured by the States. Practical considerations such as inadequately trained manpower, equipment availability, and lack of adequate end-result tests in some instances prevent an immediate, complete changeover from the *traditional* specifications. However, a number of States already are assessing the degree to which they are involved in the process control and are shifting as much of the responsibility to the contractor that is possible under present circumstances. Where adequate tests to measure finished quality are available, there is no evidence that ultimate responsibility for process quality would present a hardship to the contractor. Increased contractor responsibility coupled with proper flexibility by the State should result in better and more economical construction and provide incentive for the equipment industry to produce equipment that is capable of high-quality work as well as high production.

For certain operations, reduced-payment schedules for out-of-limits construction seems to be a necessity. The designation of really *good* material or construction and really *bad* material or construction is relatively simple. However, there is usually a grey area in which the out-of-limit material or construction may be usable, and removal and replacement operations are not warranted because of delays or other hindrances to traffic. For such material or construction the concept of partial payment is not new. In current practice, payment to the contractor is arbitrated in after-the-fact negotiations. If schedules are established before the contract is let, the contractor will be aware of the risks involved and after-the-fact penalties probably will not be necessary.

Although objections have been raised to some concepts advocated in the research program, the basic idea of adapting statistical concepts to highway construction is being well received. Research data are being used by many States to revise specification limits to allow for sampling and testing errors determined through the research studies. Only one State has progressed sufficiently to include a complete statistical approach in its standard specifications. At least five States are known to be incorporating special provisions that were calculated on a statistical basis. Five other States have written statistically based specifications for some facet of their construction, but have not used them in contractual work.

Rapid progress is being made in the adoption of control charts for displaying and analyzing data. Control charts can be used under present specifications if the inherent limitations are well understood. Their use will be greatly enhanced as more information on quality requirements and measuring techniques are developed.

Optimum use of statistics in quality assurance can come only through the adoption of end-result specifications. End-result specifications will allow the proper designation of responsibility for control and acceptance, and they are the only means through which quality measurement of a completed segment of construction will ever evolve. End-result specifications require knowledge of end requirements and must be based on measurements made on the end product. The highway industry's present inability to adequately define performance requirements and to measure performance quality dictates a major redirection of the research program.

Discussion of Research Results

Information, data, and analyses obtained through research by the States, Public Roads, and others are presented in subsequent parts of this report. These data provide support for many of the statements in this introductory section.

The indicated variation in materials and construction is, in fact, often attributable to variation in sampling and testing rather than to the materials or the construction itself. It is essential, therefore, that each State determine the sampling and testing variation associated with its current methods and personnel, and that it makes a concerted effort to reduce test variations to a minimum.

Many current specifications do not adequately allow for sampling and testing variations in the presently prescribed limits. Where such inadequacy exists, and it is impossible or uneconomical to further reduce these variations either by improving the procedure or increasing the number of tests, the specification limits should be relaxed.

When random samples are taken in sufficient number to adequately measure quality, it has been shown that a surprisingly large portion of currently acceptable construction does not comply with present limits. It may not be economically feasible to make sufficient tests for accurate measurement of quality during the control and acceptance process, but as stated earlier, the use of statistical concepts will make it possible to select the sample size in accordance with the importance of the decision being made and the economics of sampling and testing. It is therefore important that the validity of current tests as indicators of quality be studied, and that new tests be developed that will better measure the performance of the end product.

The variability of materials and construction is emphasized by the data. Present procedures usually are concerned with the average level of characteristics; however, even when the target value is met, it is shown by statistical analyses that a large portion of the materials or construction may be outside specification limits. Accordingly, variation, as well as the level of quality, should be controlled. To accomplish this control, a method of random selection of samples must be used. The adoption of random sampling by industry will significantly improve quality assurance in highway construction.

The research program has produced many other findings that will be discussed in subsequent sections. However, additional data are required before the results can be established as facts. Some of the data being received are not sufficiently complete to firmly establish the necessary basic relationships.

Discussions and findings for specific items of construction will be included in the next and subsequent issues.

REFERENCES

(1) Quality Assurance Through Process Control and Acceptance Sampling, reported by the Statistical Quality Control Task Group, Office of Research and Development, Bureau of Public Roads, April 1967.

(2) The Statistical Approach to Quality Control in Highway Construction, Research Guides, The Statistical Quality Control Group, Office of Research and Development, Bureau of Public Roads, April 1965.

Part 2—

Quality Assurance of Embankments and Base Courses

Reported by THURMUL F. McMAHON,
Principal Quality-Assurance Research Engineer,
Materials Division

Introduction

Embankments and base courses, essentially, are structural elements of the highway and are amenable to the same treatment as any other structural element with respect to design, process control, and acceptance. Their function is to provide adequate support to the pavement within the design concepts of load applications.

Density Control

The engineer has learned that proper compaction is essential to the performance properties of soil and rock material. However, the uniformity of support is as important, if not more so, than the absolute magnitude of the support offered; therefore, the control of the compaction process is one of the most important aspects in base and embankment construction.

In the 19th Century, during construction of earth dams, it was discovered that the driving of livestock, particularly sheep, across lifts of soil, as they were placed, improved uniformity of support, increased stability, and decreased permeability of the completed structure. Although many improved methods of compaction and compaction control have evolved over the years from this crude beginning, compaction control is still an item of major concern to the highway engineer.

The first attempt toward scientific control of the compaction process resulted from the work of R. R. Proctor *(1)*,[1] who developed the moisture-density relations still used in compaction specifications and control. He also developed the Proctor Needle to control the uniformity of compaction. Later the overflow-volumeter, sand-cone, and rubber-balloon methods were developed to aid in the density measurement of compacted materials. The newest, and probably the best methods of measuring moisture and density of compacted materials are those in which nuclear devices are used.

The advent of nuclear equipment not only has provided a faster and better procedure for measuring compaction but has resulted in a review of the methods and the precisions to be expected. Also, extensive studies are being made to develop better criteria than density for specifying and controlling compaction in the future.

Current practices

It has long been the custom to define desirable compaction as the degree of compaction that is above some lower limit set by engineering judgment and based on experience with various materials and performance requirements. This lower limit is described as percent of a maximum density determined in the laboratory for each type of soil to be encountered on a project.

Although most engineers have recognized that measurements of density are not absolutely reproducible in themselves, and that material variations in any embankment or base may be the rule rather than the exception, the extent of the density-measurement variations seldom has been determined. Because these variations have not been recognized, misunderstanding exists within the engineering profession

---

[1] Italic numbers in parentheses identify the reference listed on page 20.

and between engineers and nontechnical people. Engineers, as well as public agencies, have often been criticized when it was shown in subsequent test results that accepted embankments and base courses failed to meet minimum requirements even though no evidence of unsatisfactory performance existed.

A look at present specifications and compaction-measurement methods emphasizes the misunderstanding that exists. To develop measurement criteria, a series of laboratory compaction tests is run to establish the maximum density and optimum moisture content for each soil or base type. It is common practice to run one series of standard compaction tests for each material although it is fairly common knowledge that if a second series was run on another portion of the same material the results of the two tests might differ by several pounds per cubic foot. Frequently, the field technician uses density values established in the laboratory to determine percent compaction at the construction site by comparing the results of field tests with the laboratory-developed curves. He must make a judgment as to whether the type soil he has tested is the same as that for which a curve has been established. It is often apparent that his decision on which curve to use is based on density comparisons rather than on soil type comparisons. Present day construction methods further contribute to the difficulties of the technician, who seldom will encounter material in the field that is an exact duplicate of the material tested in the laboratory. Excavation and spreading of large quantities of materials nearly always result in mixtures of types or variations of type from spot to spot in the fill.

Not only are the methods of applying the test results difficult to rationalize, but the tests themselves are not reproducible to the extent necessary for exact measurement. Several years were futilely spent in comparing the results of nuclear measurements to those of conventional measurements. Only recently has it been demonstrated that the nuclear device is capable of producing more precise overall data than can be obtained by conventional methods.

One major factor that influences the variation in conventional-density test results is the common practice of removing the larger particles, greater than ¾ inch, from the samples tested in the laboratory. The effect of these larger particles on field results is estimated by empirical mathematical equations and superimposed on the results of the laboratory tests. Many laboratories realize the fallacy of this practice and are using larger molds in their tests.

Sampling

Selective sampling by the inspector, often as ordered by the engineer, has played an important part in the failure to recognize the magnitude of the actual variations occurring in embankment and base construction. When the inspector has the opportunity to select the test site, he has three alternatives: (1) To select an average condition, (2) to select the poorest condition, and (3) to select the best condition. The general custom in the State or the specific practice of the engineer on the job may well determine the site he selects for test. Regardless of his choice, the results of his tests will reflect only the condition he is selecting and not the variability of results or the true overall level of compaction.

Valid measurements of the actual quality of the compaction can be made only if the sample is a true representation of the total compacted material. It is possible to obtain a representation of the entire mass only when the sampling program is so designed that each element in the mass has an equal chance of being one of the elements of the sample. Of course, the greater the number of elements sampled, the better will be the representation.

The Statistical Approach

Although many questions concerning the required level of compaction and the methods of obtaining it are still unanswered, almost everyone agrees that uniformity of support is the principal requirement of good embankment and base-course construction. As a result of recent measurements obtained in research, the need for a change in methods of control has become apparent. Any such change must be directed toward controlling uniformity as well as degree of compaction.

The use of statistical concepts to establish the requirements of specifications and to aid in the analysis of test data provides much of the needed improvement. The specification either designates a target percent-compaction value and the allowable variations about this value or designates a lower limit to be met by a given percentage of the construction, when a valid statistical analysis of test results is performed.

A statistically based specification requires that a contractor submit a *lot* of predetermined size to the buyer for acceptance. Each *lot* is evaluated on the basis of the results of a specified sampling and testing program. This program entails the performance of a specified number of standard tests at random locations on each *lot* submitted. The data analysis procedure to establish compliance and the steps to be taken if noncompliance is indicated are also spelled out.

Several States have developed specifications for embankment or base construction that are great improvements over present methods and are based partly on statistics, even though they are not strictly in accordance with concepts recommended in this series of discussions.

Virginia, for example, is using a control strip technique for control of the compaction of aggregate base. The following special provisions were extracted from a paper *(2)*, presented at the 46th Annual Meeting of the Highway Research Board:

"Virginia Department of Highways Special Provisions For Nuclear Field Density Testing of Aggregate Base and Surface Courses

"Section 308 of the 1966 edition of the Road and Bridge

Specifications is amended in this contract to require the construction of density control strips for the purpose of using the nuclear field density testing device. The revisions are as follows:

"At the beginning of the work the Contractor shall build a control strip of the material on an approved and stable subgrade for the purpose of the Engineer's determining density requirements for the project. This control strip will be at least 400 square yards in area and of the same material and depth to be used in the remainder of the work. Compaction will be carried out with conventional rollers approved by the Engineer until no appreciable increase in density is accomplished or until in the opinion of the Engineer no appreciable increase in density will be obtained by additional rolling. Upon completion of the rolling, the density of the strip will be determined by use of a portable nuclear test device.

"The compaction of the remainder of the aggregate base course material shall be governed by the density of the control strip. The material shall be tested by sections of approximately 2,800 square yards each. The mean density of 5 randomly selected sites from the test section shall be at least 98 percent of the mean density of 10 tests taken from the approved control strip. Placing, compacting and individual testing may be done in subsections of approximately 280 square yards each. When the mean of the test section is less than 98 percent of the control strip mean the Contractor may be required to rework the entire section. Also, each individual test value shall be at least 95 percent of the mean value of the control strip. When an individual test value is less than 95 percent of the control strip mean, the contractor shall be required to rework the area represented by that test.

"Each test section shall be tested for thickness and any deficiency outside the allowable tolerance shall be corrected by scarifying, placing additional material, remixing, reshaping and recompacting to the specified density.

"A new control strip may be requested when:

(1) A change in the source of the material is made, or
(2) a change in the material from the same source is observed, or
(3) ten (10) test sections have been approved without the construction of additional control strips.

"Note: The Contractors' attention is directed to the fact that the method for determining density and the requirements for density as described in Section 308.05 have been replaced by the method of determination and requirements for density stated hereinabove."

Reliability of nuclear testing

The Virginia specification is an excellent example of the more rapid methods that can be used to allow the testing of a more representative sample of completed work. The increased number of test results available for making a decision assures a higher confidence in the decision. The following advantages are claimed by the Virginia Department of Highways:

- Nuclear tests can be made quickly and easily.
- A field control strip provides a practical achievable density.
- The speed of nuclear testing permits determination to be made for each section of material. This procedure provides a sound statistical basis for decisionmaking.

The reliability of nuclear-gage test results is substantiated by tests made in a number of States. For example, the data in table 1 are from two studies in Utah (*3*, *4*). The absolute values of the standard deviations presented in the table have little significance with respect to testing variability because much of the indicated variation is probably caused by actual density variations. However, it is significant that the sampling and testing variance is smaller and that there is no significant difference in the means. The results of the nuclear tests are as good, if not better, than those of the conventional tests; consequently, it can be stated safely that the testing error of nuclear methods is probably no greater than that of conventional methods.

Reported variations in compaction

The variation in density of accepted embankments and bases has been found much greater than had been expected when the Public Roads research program (see part 1 of this article, Feb. 1969 issue) was initiated. Because of this variability, compliance with specifications, as computed by statistical methods, is lower than had been expected. Therefore, designers must judge whether present construction is sufficient for their purposes. If present construction is satisfactory, then specifications should be changed to allow for the existing variation. If better construction is needed, then it is important that specifications and methods be changed to assure better uniformity in embankments and base courses.

Research is showing that overall standard deviation, a measure of variability, is not in itself a true indication of contractor-performance variability. A good contractor may take the same care in constructing two embankments but the

Table 1.—Percent relative compaction for different test methods

Compacted components	Sand cone			Portable nuclear			Roadlogger		
	Mean	[1]σ_{st}^2	[2]σ_o	Mean	[1]σ_{st}^2	[2]σ_o	Mean	[1]σ_{st}^2	[2]σ_o
Embankments	99.1	3.46	4.46	99.0	2.25	4.55	100.2	2.17	4.44
Bases and subbases ...	98.7	2.16	2.92	98.1	1.32	2.89	98.0	0.60	2.48

[1] σ_{st}^2 Sampling and testing variance.

[2] σ_o Overall standard deviation.

variability of test results may be much greater in one than in the other. If the composition of the material itself is more variable, then the results of the compaction process are also going to be more variable.

The variation of density in embankments, with respect to material and process changes, is shown in figures 1 and 2. Figure 1, extracted from a California report (5), presents the distributions of the results of density tests on three projects.

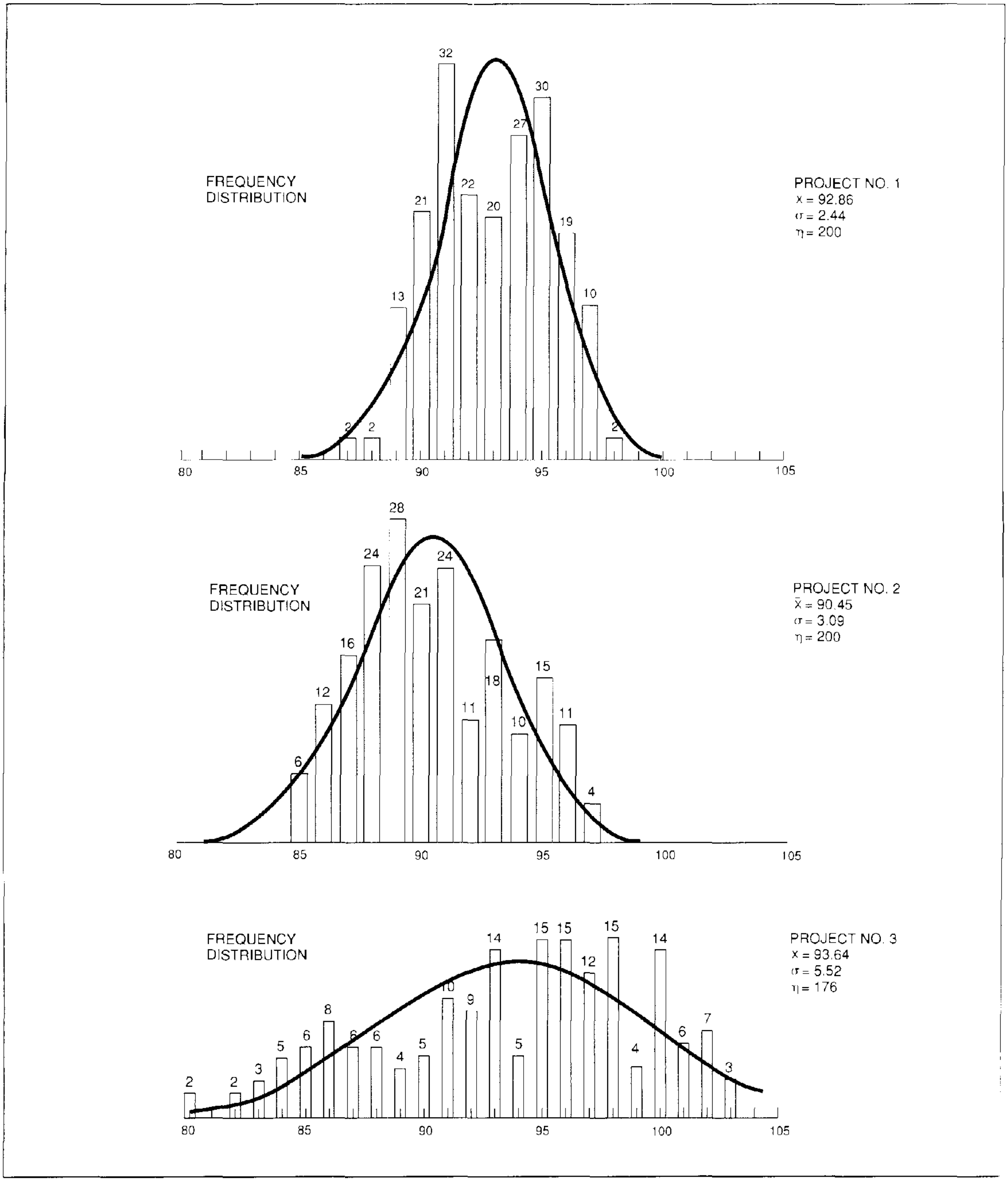


Figure 1.—Variation in density of embankments, California road-embankment study.

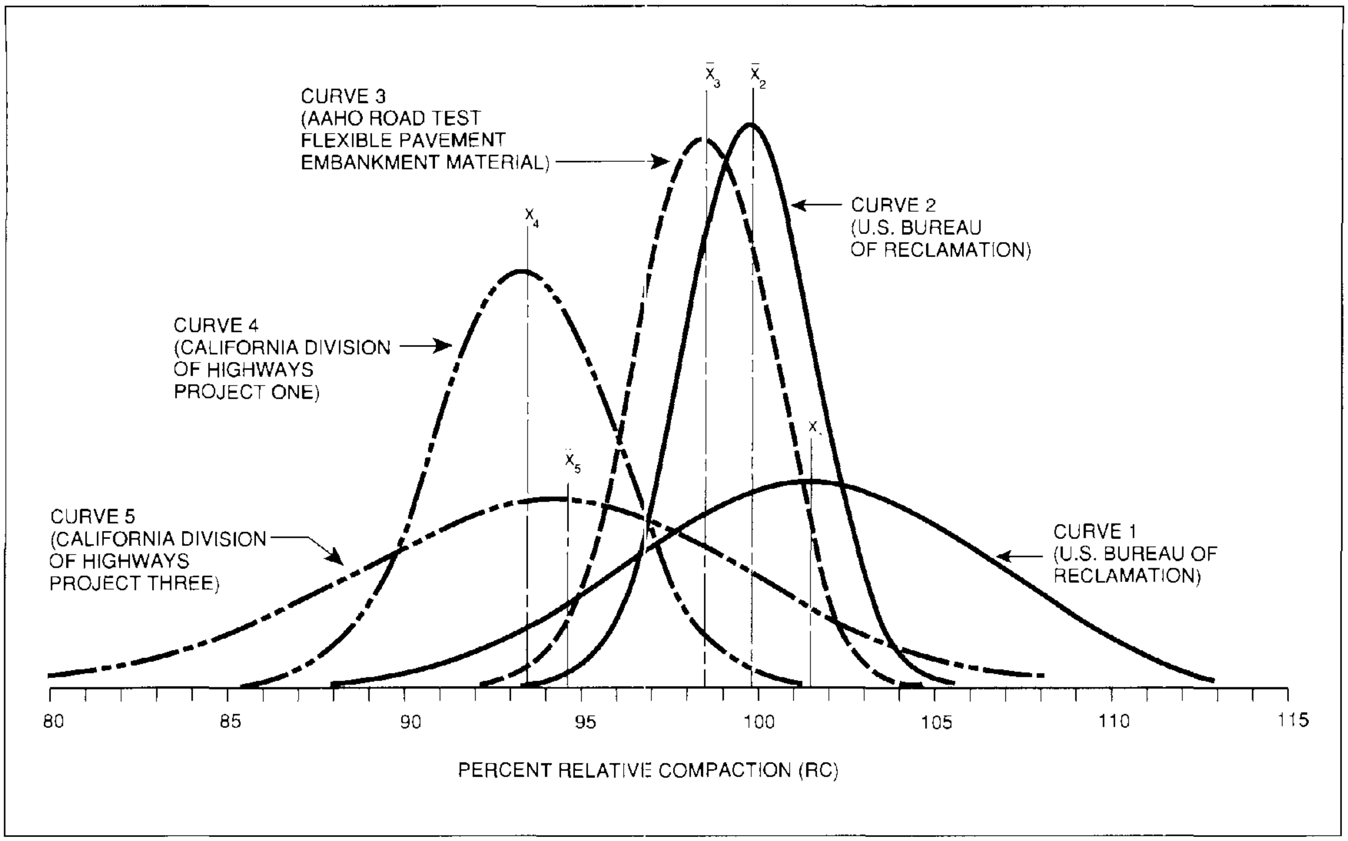


Figure 2.—Normal distribution curves from three organizations.

Project No. 1 was constructed with homogeneous, fine grained soils; Project No. 3 with an extremely heterogeneous soil; and Project No. 2 with a soil of intermediate variability, with respect to the other two. The specification on each of the projects stipulated that the material be compacted to no less than 90 percent relative compaction. It has been shown by many of the research test results obtained after acceptance by normal control procedures, that the construction does not meet specification requirements when the data are analyzed on a statistical basis and the total material is considered.

Figure 2, also from the California report, is presented to show that variability of compaction test results is not unique to the highway industry.

Figures 3 and 4 have been extracted from an Alabama Research Report (6) to show indicated variation in density of compacted base and subbase materials. The standard deviations of 4.06 and 2.31 percent are in line with values reported by other States. Figures 5 and 6 are from the same report; the data reported by California for compacted densities of embankment materials are corroborated by the data in figure 5, and variations in moisture content are shown in figure 6. The large variation in moisture content is probably a major cause of the large variation in density.

From a research study performed by Purdue University for the State of Indiana in March 1967 (7), information concerning average density, range, and standard deviation are shown in table 2 for three subgrade and three subbase projects. The data for the study were obtained after the projects had been accepted under normal acceptance procedures. The specifications for the projects required a minimum density of 100 percent of standard laboratory maximum density.

The wide ranges of results and large standard deviations reported in table 2 are, in part, due to variability contributed by test methods. The differences between replicate sand cone density tests on the study projects are shown in table 3. The entire difference cannot be attributed to test error as there may be actual differences in the materials or densities even when the tests are taken side by side as was done in this study. However, the results show the magnitude of differences when an effort was made to eliminate material differences within the limitations of practical construction conditions.

Another contributing factor to the variation of test results is the difference in results of laboratory maximum-density and optimum-moisture-contest tests. These differences for duplicate samples from Project S-1 of the Purdue University study are shown in table 4.

The data obtained in the Purdue University study are shown in figures 7 and 8 which are histograms of the percent compaction for the six projects.

Figures 9 and 10 were extracted from a report of a study conducted by the Engineering Experiment Station of North Dakota State University for the North Dakota State Highway

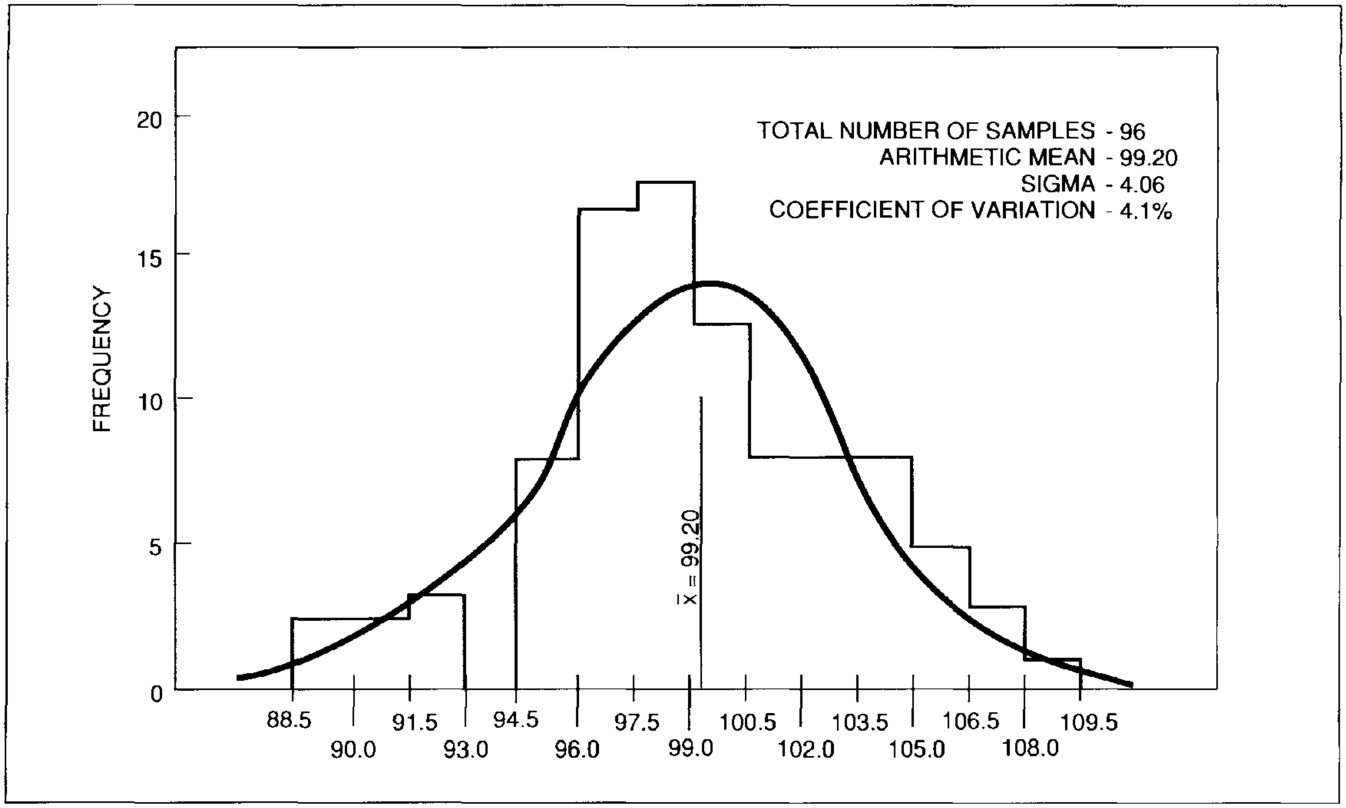


Figure 3.—Soil aggregate base, percent compaction.

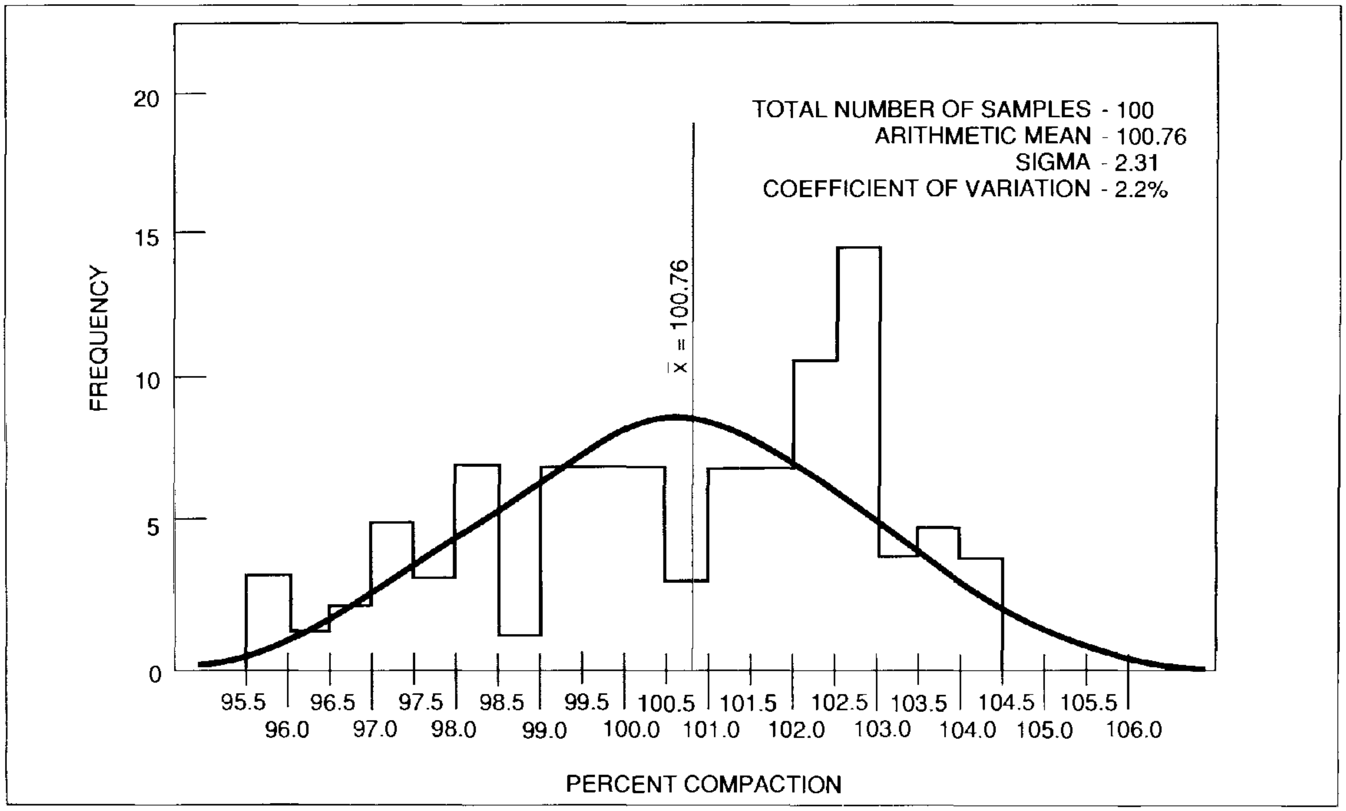


Figure 4.—Selected soil subbase, Class 4, percent compaction.

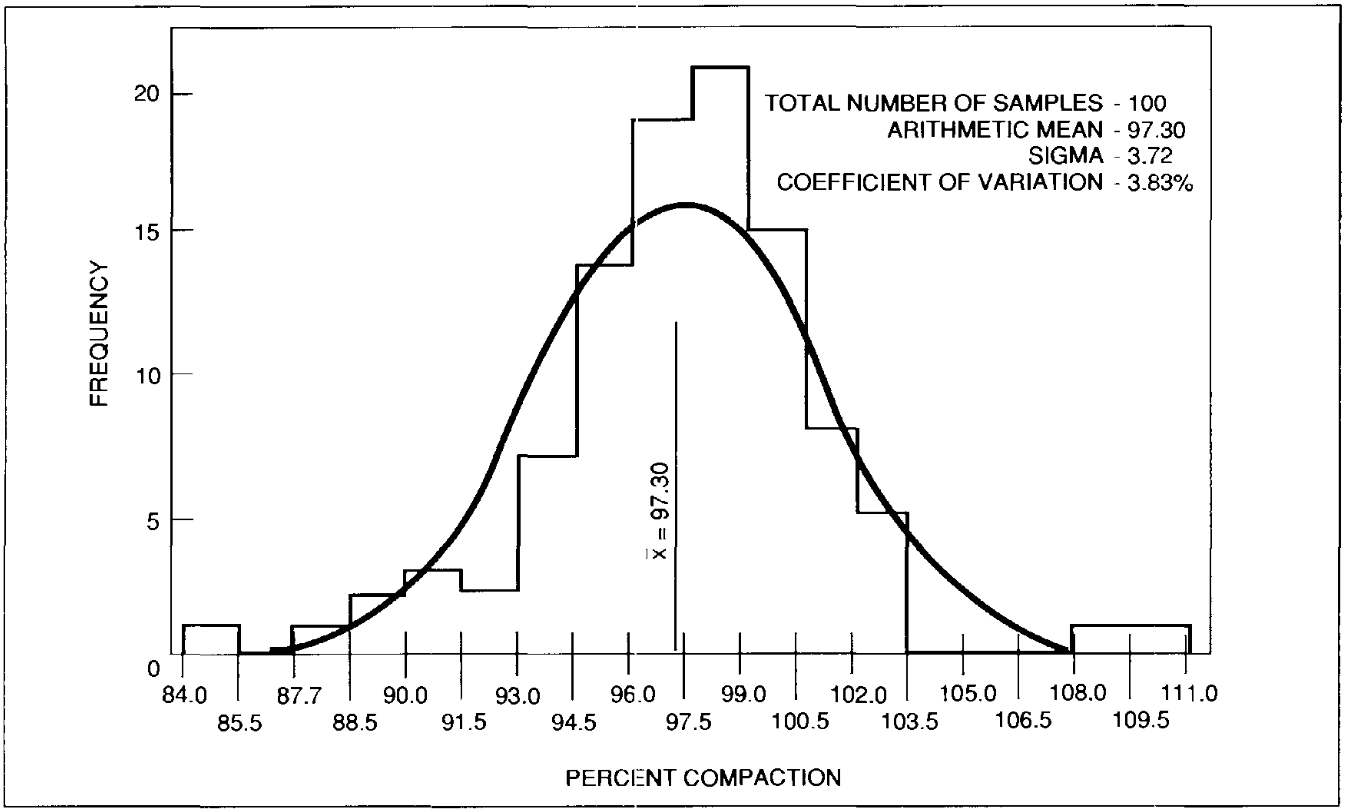


Figure 5.—Embankment, percent compaction.

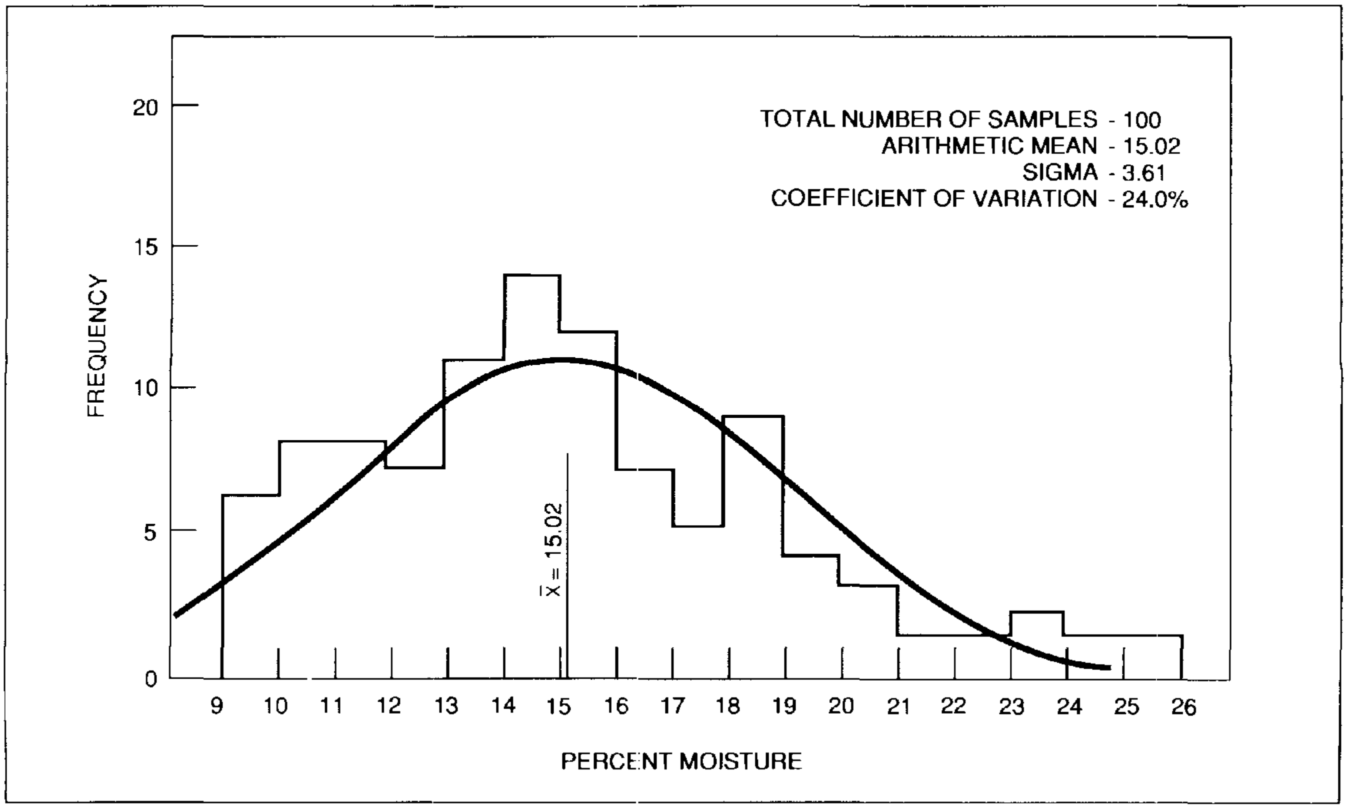


Figure 6.—Embankment, percent moisture.

Table 2.—Average, range, and standard deviations of percent compaction of subgrade and subbase projects

Project	Average compaction	Range of compaction	Standard deviation
	Percent	*Percent*	*Percent*
S–1	100.6	84–116	5.3
S–2	96.8	80–110	5.7
S–3	98.2	84–108	4.5
B–1	89.4	82–98	3.3
B–2	91.7	84–100	3.1
B–3	93.6	86–100	2.3

Table 3.—Average differences between sand-cone density tests for replicate tests

Project	Replicates	Average difference between sand-cone density values for replicate tests
	Number	*Lb. per cu. ft.*
S–1	48	3.32
S–2	48	4.95
S–3	49	4.18
B 1	51	4.15
B–2	55	3.35
B–3	50	2.24

Table 4.—Maximum dry density and optimum moisture content values for duplicate field samples for Project S–1

Test number	Maximum dry density	Optimum moisture content	Liquid limit	Plasticity index	AASHO classification
	Lb. per cu. ft.	*Percent*			
27 A	118.7	13.7	32.5	13.5	A–6 (8)
27 B	119.3	13.5	36.0	15.2	A–6 (9)
32 A	117.4	14.4	33.9	14.3	A–6 (6)
32 B	115.8	15.4	34.6	13.9	A–6 (8)
34 A	113.5	15.4	41.5	18.4	A–7–6 (11.5)
34 B	110.6	14.3	41.6	16.9	A–7–6 (9)
36 A	123.0	13.4	30.2	12.3	A–6 (7)
36 B	119.7	13.3	31.0	11.9	A–6 (8)
37 A	121.2	12.6	28.1	8.6	A–4
37 B	122.6	11.0	—	—	—
38 A	112.6	17.3	41.3	17.6	A–7–6 (11)
38 B	113.5	15.9	37.6	14.6	A–6 (9)

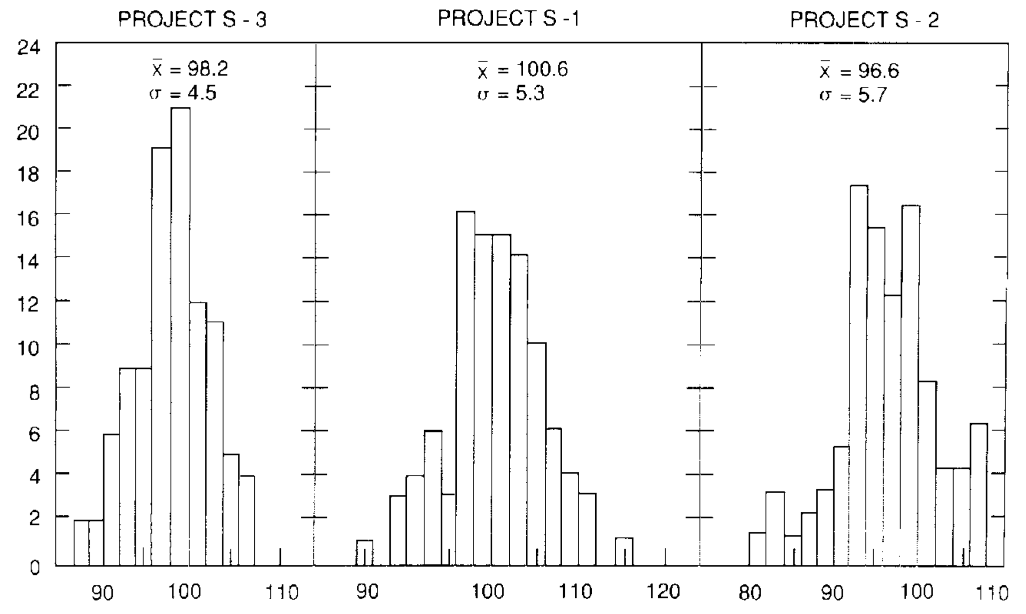


Figure 7.—Frequency histograms—percent compaction of subgrade materials for three projects.

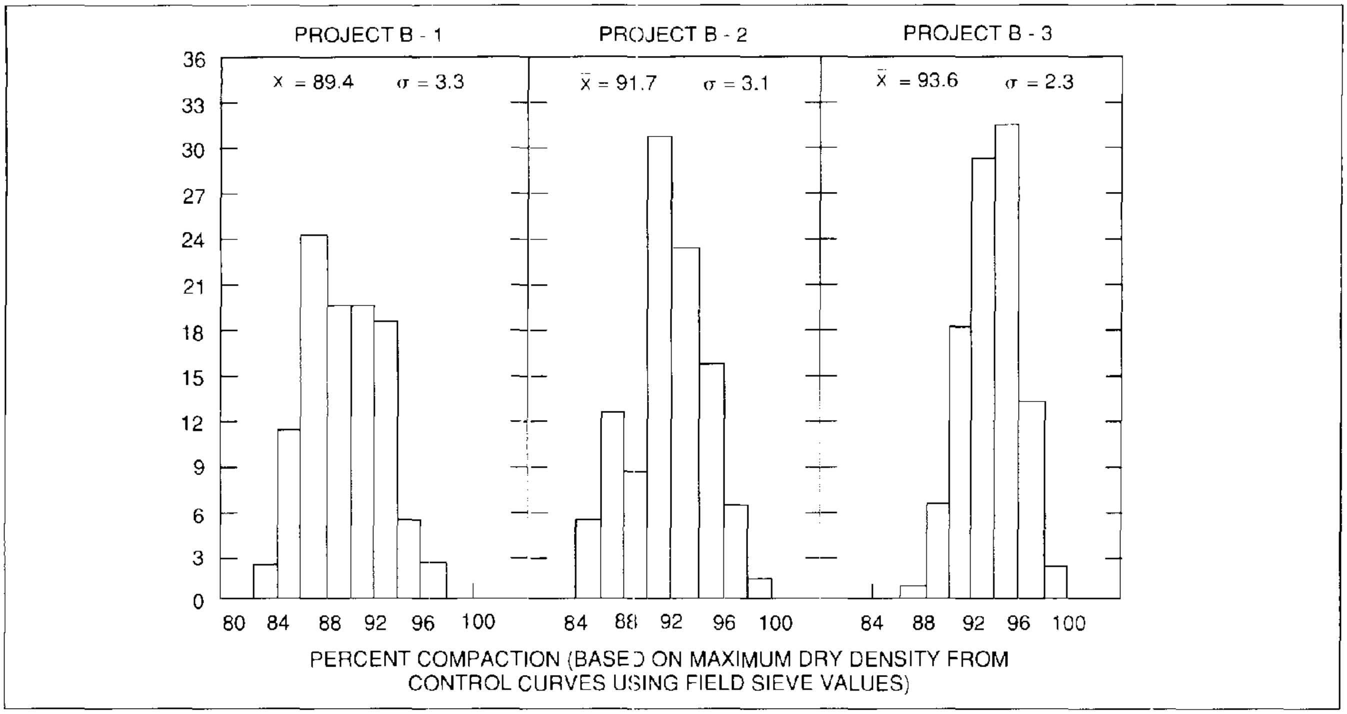


Figure 8.—Frequency histograms—percent compaction of subbase materials for three projects.

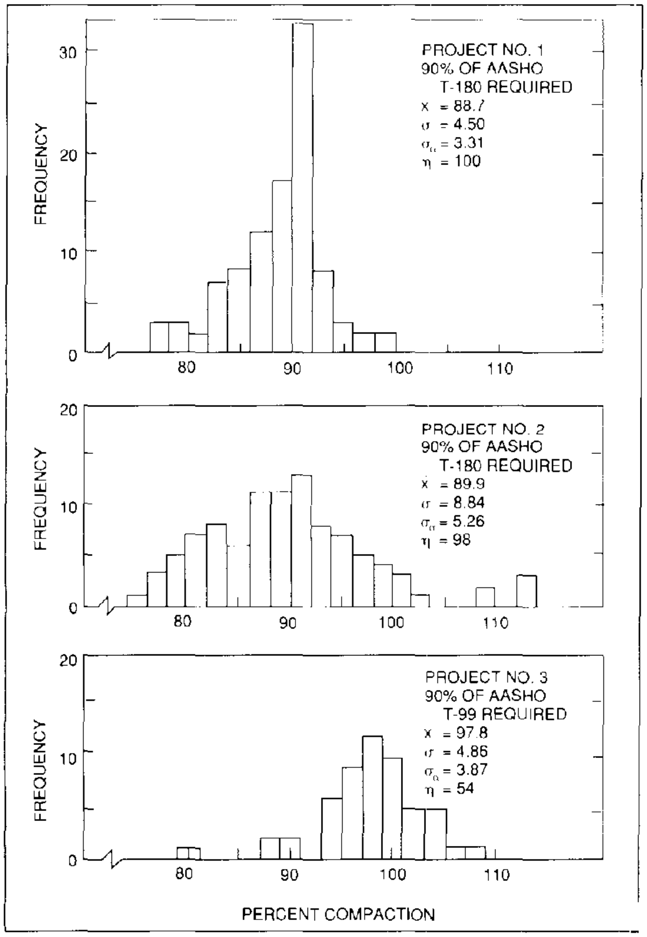


Figure 9.—Percent compaction—research data.

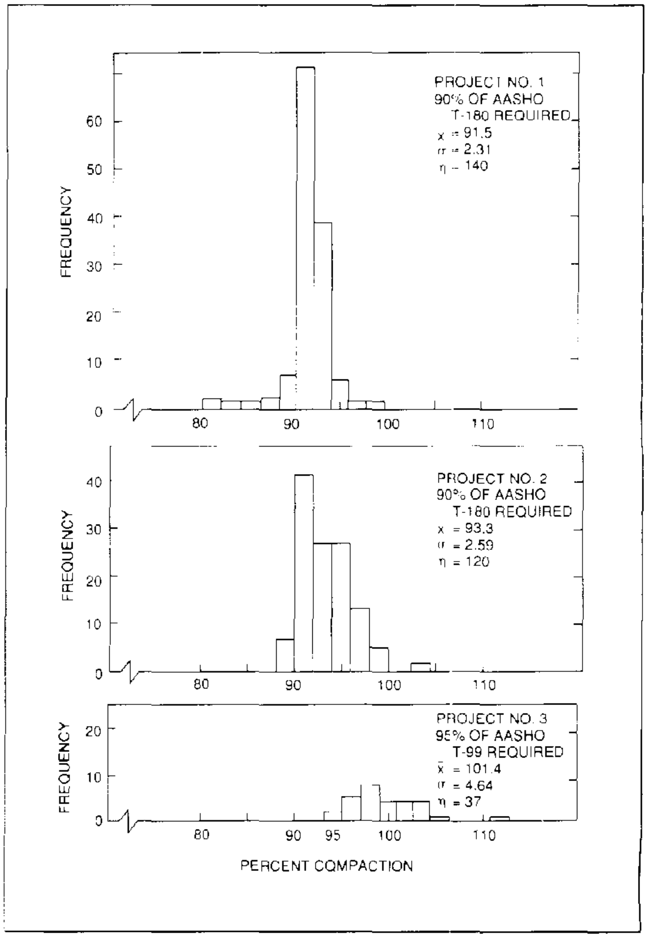


Figure 10.—Percent compaction—highway-department data.

Department (*8*). In figure 9 is shown the variability of compaction in three embankment projects previously accepted under current control and acceptance procedures. The distribution of test results on random samples are presented in the histograms in the figure. The mean $\bar{X}$; the overall standard deviation, σ; and the sampling and testing standard deviation, σ_α, of the distributions are tabulated. These standard deviations must be changed to variances in order to obtain the relationship between the material and the sampling and testing variability, $\sigma^2 - \sigma_\alpha{}^2 = \sigma^2{}_m$. In figure 10, the information obtained during routine control and acceptance testing on the same three projects is presented. Comparison of the results presented in figures 9 and 10 emphasizes the advantages of random sampling in determining the true as-built conditions of any construction project and compliance with specifications.

Density-test results obtained with two types of nuclear gages and two different test methods on Project No. 1 of the North Dakota study are shown in figure 11. These data substantiate the results of the Utah report in that the sampling and testing errors for the nuclear devices are smaller than those for the water-balloon method (fig. 9). It is of interest that the air-gap method indicates a higher average density than the water-balloon or contact-nuclear method. A similar nuclear study

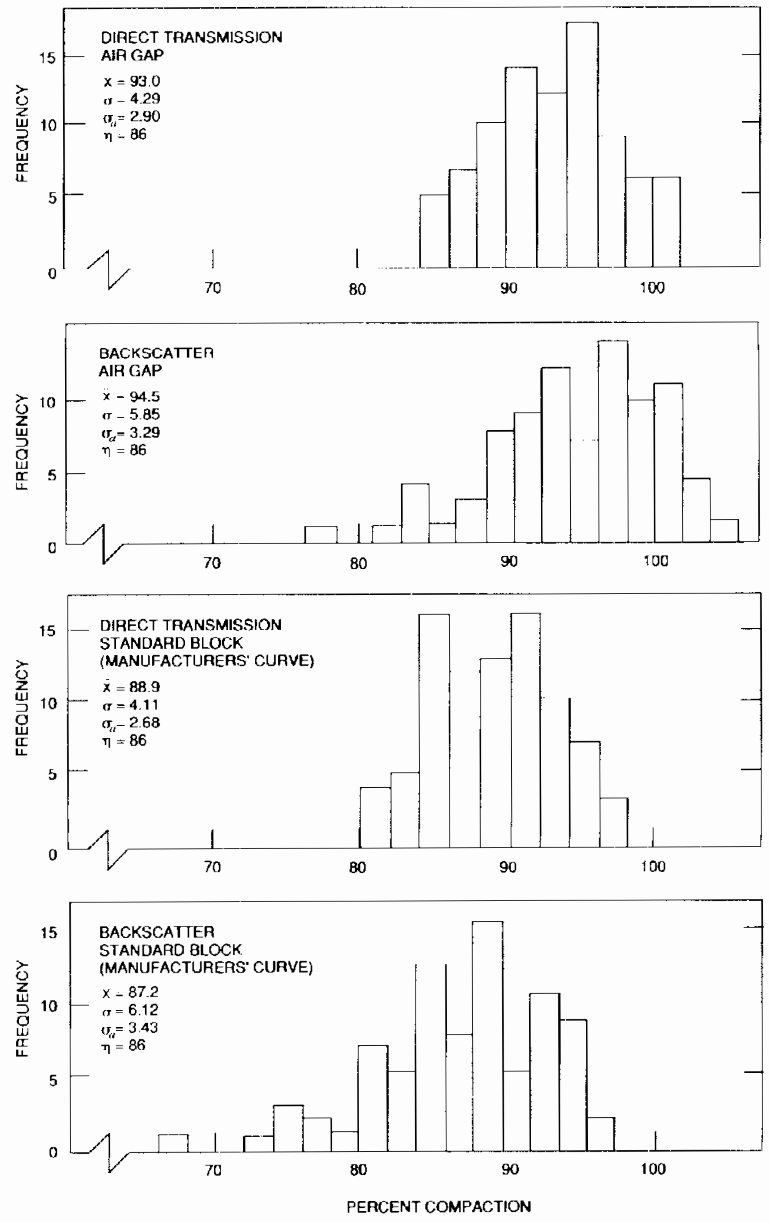


Figure 11.—Percent compaction—nuclear-instrument data.

was performed on Project 2 with parallel results. In these tests, the manufacturer's calibration curves for the nuclear devices were verified before use.

Variations in Material Properties

Tables 5 and 6 were extracted from a California report (*9*) to show the variation of test results other than those of density tests. The data are from six projects selected as typical of material used for untreated base and subbase by the California Division of Highways. Again the data were obtained from random samples taken after the materials had been accepted as complying with the specifications for normal sampling methods. These materials were largely in substantial compliance with the specifications; however, there was considerable variation in the test results of the material properties, which may account for some of the variations in density and supporting capacity exhibited by the compacted material. The study did not include the determination of density variation of the in-place material.

Conclusions

The primary conclusion that can be drawn from the data presented in this discussion is that test results on base and embankment materials exhibit significant variation. These variations can be attributed to material variance, sampling variance, and testing variance. Many materials may be classified out-of-specification because of sampling and testing errors rather than failure of the material or construction to actually conform to specified requirements.

It should be apparent that improvement of sampling and testing methods must be a priority research and development item if field measurements on samples are to be used to accept construction materials and structures. More tests results must be used in the decision process to increase the validity of decisions. Rapid sampling and testing methods, together with random sampling and statistically valid decision plans, will alleviate many of the problems in current acceptance of construction.

The data and charts of this presentation clearly indicate a difference between the test results on random samples and those on representative samples. A true estimate of the actual quality of any material or construction can be obtained only when every item of the *lot* has a chance of being chosen as part of the sample. Sampling by choice cannot provide samples that will permit evaluation of both level and variability of material or construction. Randomizing sampling locations is a simple matter and should cause no serious problems for the inspector, especially when rapid nondestructive test methods,

Table 5.—Summary of test results, untreated aggregate base

Test	[1] n	[2] $\bar{X}$	[3] σ	Range of results	Specification requirement	Amount not complying with specification
PROJECT B-1						
						Percent
R value	200	81.9	1.3	78–85	[4] 78	0
Sand equivalent	200	42.9	4.0	33–58	[4] 30	0
Percent passing #4 sieve	200	50.9	3.1	35–55	35–55	0
Percent passing #30 sieve	200	23.8	2.5	15–30	10–30	0
Percent passing #200 sieve	200	6.0	0.7	4–8	3–9	0
PROJECT B-2						
R value	200	79.9	2.4	72–85	[4] 75	2.5
Sand equivalent	200	30.6	6.1	24–63	[4] 30	56
Percent passing #4 sieve	200	58.1	2.8	51–67	35–65	2.5
Percent passing #30 sieve	200	27.3	2.3	22–36	(5)	—
Percent passing #200 sieve	200	7.9	1.1	4–10	3–12	0
PROJECT B-3						
R value	200	79.7	1.5	78–83	[4] 78	0
Sand equivalent	200	59.2	4.0	48–68	[4] 30	0
Percent passing #4 sieve	200	52.7	5.7	40–71	35–55	33
Percent passing #30 sieve	200	23.4	2.9	15–31	10–30	0.5
Percent passing #200 sieve	200	4.6	0.95	3–12	3–9	0.5

[1] n = Number of samples.
[2] $\bar{X}$ – Arithmetic mean.
[3] σ = Standard deviation.
[4] Minimum.
[5] None.

such as the nuclear gage, are available. State highway departments should take immediate steps to implement random sampling in the control and acceptance of base and embankment construction.

The data reported here concerning the variations in base and embankment construction should not be taken as an indictment of present construction. Although there is adequate information to indicate that improvement is needed in the testing and analysis of data, there is no specific information available to indicate that construction being accepted under present procedures is not performing to design expectations. However, if economic considerations do not allow an intensive effort to reduce sampling and testing variation, as well as actual variation in density and moisture content, it is imperative that recognition be given the variations occurring in present construction and that current specifications be revised accordingly.

REFERENCES

(*1*) *Fundamental Principles of Soil Compaction*, by R. R. Proctor, Engineering News Record, VIII, vol. Nos. 9, 10, 12, 13, Aug. 31, Sept. 7, 21, and 28, 1933.

(*2*) *Compaction Control of Granular Base Course Materials by Use of Nuclear Devices and a Control Strip Technique*, by M. C. Anday and C. S. Hughes, Highway Research Record Number 177, Symposium on Compaction of Earthwork and Granular Bases, 1967, pp. 136–143.

(*3*) *Characteristics of Compacted Embankments*, by Frank C. Van Houten, Utah State Highway Department, August 1967.

(*4*) *Characteristics of Compacted Bases and Subbases*, by Gary F. Nielson, Utah State Highway Department, August 1967.

(*5*) *A Statistical Analysis of Embankment Compaction*, by Geo. B. Sherman, Robert O. Watkins, and Rogel Prysock, State of California, Department of Public Works, Division of Highways, May 1966.

(*6*) *Quality Control of Construction by Statistical Tolerances*, by J. H. David, Alabama Highway Department, May 1967.

(*7*) *An Investigation of Compaction Variability for Selected Highway Projects in Indiana*, by T. G. Williamson and E. J. Yoder, Purdue University, Indiana State Highway Commission, December 1967.

(*8*) *The Statistical Approach to Quality Control in Highway Construction*, by Dr. James L. Jorgenson, North Dakota State University, No. 15, Engineering Experiment Station Series, November 1968.

(*9*) *A Statistical Analysis of Untreated Base and Subbase Material*, State of California, Department of Public Works, Division of Highways, March 1967.

Table 6.—Summary of test results, untreated aggregate subbase

Test	[1] n	[2] $\bar{X}$	[3] σ	Range of results	Specification requirement	Amount not complying with specification
PROJECT S–1						
						Percent
R value	200	68.8	6.4	47 80	[4] 60	9.5
Sand equivalent	200	30.2	4.0	22–46	[4] 25	5.0
Percent passing #4 sieve	200	49.5	4.3	40–60	35–65	0
Percent passing #200 sieve	200	7.8	1.3	5–13	3–11	0.5
PROJECT S–2						
R value	188	77.2	3.1	66–83	[4] 60	0
Sand equivalent	188	36.2	8.5	14–69	[4] 25	2.0
Percent passing #4 sieve	188	72.6	6.5	60–91	30–100	0
Percent passing #200 sieve	188	10.0	1.8	5–16	0–20	0
PROJECT S–3						
R value	200	70.9	8.9	42–84	[4] 55	6.5
Sand equivalent	200	29.2	2.7	21–37	[4] 25	4.5
Percent passing #4 sieve	200	45.0	6.6	23–57	35–80	4.5
Percent passing #200 sieve	200	8.6	1.7	4–12	5–35	1.0

[1] n = Number of samples.
[2] $\bar{X}$ = Arithmetic mean.
[3] σ = Standard deviation.
[4] Minimum.

Part 3–
Quality Assurance of Portland Cement Concrete

Reported by WESLEY W. BAKER, Highway Research Engineer, and THURMUL F. McMAHON, Principal Quality-Assurance Research Engineer, Materials Division

Introduction

Ever since the development of portland cement concrete for use in the construction of highway pavements and structures, highway engineers have been concerned with the quality of the concrete and of its constituents. From years of experience, methods have been developed to control the quality of concrete and measure its acceptability as a quality material. But responsibility for the quality of portland cement concrete and causes of its failure are still confusing issues.

Research has provided an insight into the causes of variations that have always existed in the results of concrete tests. It has also provided some knowledge of their magnitude as regards blending, mixing, sampling, and testing, especially under laboratory conditions. In early research it was recognized that statistical concepts afforded a useful tool to analyze the data and establish the causes of variation. In fact, as early as the 1940's, Mr. Alfred M. Fruedenthal recommended that statistical concepts be used to revise portland cement concrete specifications (*1*).[1]

A committee on quality control of concrete in the field, appointed in England at the beginning of the 1950's, reviewed all aspects of concrete production and recommended methods of improving quality and testing techniques. The committee's greatest achievement was the adoption of statistical concepts to better understand the nature of variation in concrete production. The normal distribution was shown to be applicable in the concrete industry, particularly for the cylinder strength distribution, and acceptance criteria were established at a 95-percent confidence level.

The first official action in the United States to adapt the statistical approach to quality control of concrete was taken in 1955 by the American Concrete Institute. Criteria were established and rational specifications for structural concrete were recommended.

In this report on the application of statistical quality control methods to the production of concrete by the highway industry (*2*), Mr. Edward A. Abdun-Nur advocated extensive use of statistical concepts in specification writing and acceptance sampling. He considered a realistic picture of concrete produced under normal control to be one in which the coefficient of variation of 28-day strength is 20–25 percent, and defined a good concrete as one with a 15-percent coefficient of variation.

Early work in quality control of portland cement concrete demonstrated the advantages of using statistical concepts to specify, control, and accept concrete; however, more information was needed to make optimum use of them. The Office of Research and Development, Bureau of Public Roads, has been promoting the gathering of information by the States concerning the quality of the concrete being produced under current specifications and the contributing factors in the variation of test results. The objective of this research program has not been to determine all factors relating to variations in concrete production, but to isolate variations owing to materials, sampling, and testing. The results to date are presented and analyzed in the following paragraphs.

Variability in Concrete Strengths

Strength is not always the most important characteristic of concrete quality, but it is the one that is most often measured. It is assumed to be indicative of the water-cement ratio and, accordingly, an indicator of durability. The magnitude of the variability in strength is, therefore, an indicator of the magnitude of variability of the other characteristics.

---

[1] Italic numbers in parentheses identify the references listed on p. 29.

Variability in concrete strengths can be attributed to two other types of variability: (1) Inherent variability in the materials and processes that results from chance causes, which cannot be controlled, and (2) variability from assignable causes which can be controlled. The attainable quality of concrete in the field is limited not only by the chance causes that contribute to the variation in quality, but also by the economic factors entailed in reducing the assignable causes.

The measurement of variability from chance or inherent causes is complicated by the inherent variability of each of the ingredients in the mix, which can interact with the processes of blending, mixing, and placing, and result in a much larger variability in the concrete itself.

The assignable causes of variability are more numerous and more difficult to isolate, but the production of quality concrete is dependent on the reduction of all variables. However, the isolation and restriction of variables can be carried only as far as economic conditions warrant. Under the present state of knowledge, the ultimate uniformity of concrete production cannot be precisely stated. Extensive research will be necessary to isolate variables and to determine the extent to which variation can be reduced. Current information can be used only to show that variation does exist, and that sampling and testing often contribute as much of the variability as do the ingredients and processes used in construction.

Mr. H. H. Newlon, in a paper (*3*) described and discussed numerous variables affecting concrete quality. The data concerning concrete variability, presented in table 1, was based on a similar tabulation from his paper. Although such information is interesting and may be used to design specification limits, it is of little worth to the overall problem of reducing variability in concrete construction. The basic need is to isolate the common factors affecting variability. Research aimed at this purpose has been underway for the past 4 years.

Data based on a West Virginia research report (*4*) are presented in table 2. These data are illustrated in figure 1, which depicts the relations among the materials, sampling,

Table 1.—Average deviation of concrete strengths

Agency	Concrete type	Data source	Average standard deviation [1] $(\bar{\sigma})$
			p.s.i.
Bureau of Public Roads	Paving	Research	585
Do	.....do	Historical	473
Do	Structural	.....do	576
Virginia Department of Highways	.....do	Special	467
Do	Paving	Cores	663
Ontario Department of Highways	.....do	Routine	494

[1] The average standard deviation for all data presented.

Table 2.—Portland cement concrete variations

Project No.	28-day compressive strength variations								
	Mean $(\bar{X})$	Overall standard deviation (σ_o)	Overall coefficient of variation (υ_o)	Standard deviation, testing (σ_t)	Coefficient of variation, testing (υ_t)	Standard deviation, sampling (σ_s)	Coefficient of variation, sampling (υ_s)	Standard deviation, materials (σ_a)	Coefficient of variation, materials (υ_a)
Structural concrete									
	p.s.i.	*p.s.i.*	*Pct.*	*p.s.i.*	*Pct.*	*p.s.i.*	*Pct.*	*p.s.i.*	*Pct.*
1	4,235	435	10.0	170	4.0	236	5.6	310	7.2
2	4,420	482	10.9	323	7.3	39	0	360	8.1
Paving concrete									
1	4,675	545	11.7	377	8.1	91	0	386	8.3
2	3,755	420	11.2	322	8.5	42	0	270	7.1
3	3,720	575	15.5	318	8.5	....	0	495	13.3
4	4,760	467	9.8	200	4.2	34	0	420	8.8
5	4,688	733	16.5	585	12.5	....	0	545	11.7

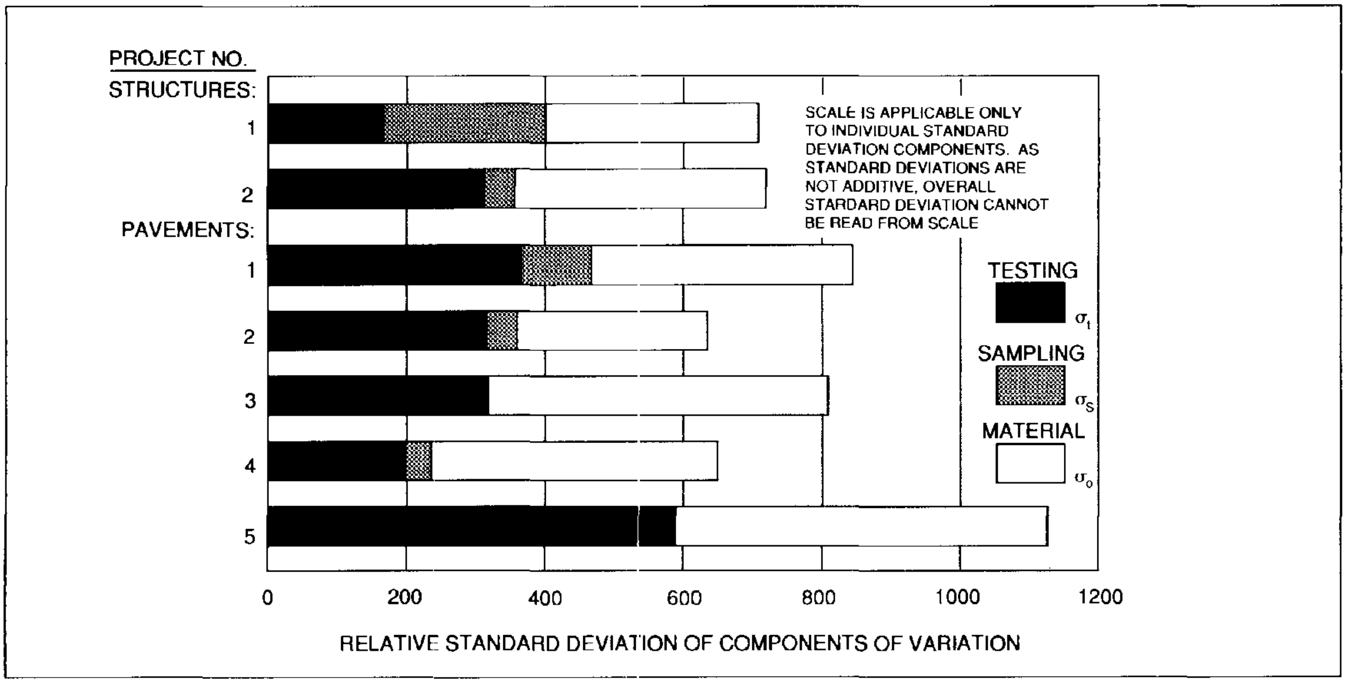


Figure 1.—Portland cement concrete—standard deviation, compressive strength.

and testing standard deviations. As standard deviations are not additive, the sum of the standard deviation shown does not equal the standard deviation of the concrete strength. A significant indication from these data is that the combined testing and sampling variations are usually greater than the material variation. It is also significant that the materials deviations consist of the material and process variations, whereas the sampling and testing deviations are caused by the measurement process. Figure 1 indicates that sampling did not contribute significantly to the variation of test results on these projects.

An analysis of historical data on compressive strengths of concrete cylinders, presented in table 3, was based on a report by the State Road Department of Florida (*5*). Two types of concrete, class *A* and class *NS* were analyzed in the report. Routine control is normally exercised over class *A* concrete, whereas class *NS* concrete is spot checked only occasionally.

Based on the same Florida report the mean strength, standard deviation, and coefficient of variation for the concrete, shown in table 3, were compared by source in table 4. Researchers have shown that the standard deviations of strength test results usually increase with the mean or average strength of the concrete; therefore, the best comparison of these data may be shown by the coefficient of variation. As expected, the coefficients of variation for class *NS* concrete were greater than those for class *A*.

The strength data presented—typical of nearly all strength data received—expressed large standard deviations with average strengths well above the usual minimum of 3,000 p.s.i. For example, if the class *A* concrete of project 4, table 3, were analyzed using normal distribution methods, there would be less than 1 percent chance that a test would result in a compressive strength of less than 3,988 p.s.i. With the same standard deviation, the mean could be as low as 4,698 p.s.i. before a 1-percent chance of being below 3,000 p.s.i. was exceeded. If the standard deviation could be reduced, the mean might be lowered further without risking noncompliance. However, as pointed out previously, the durability of the concrete may be the controlling factor in reducing strength through lowering the design cement content. If it is assumed that strength is an indicator of water-cement ratio, it is possible that more uniform strength will also result in a more uniformly durable concrete. This in turn may allow a reduction of the design cement content.

If durability should be the controlling factor and if it should have a relation with strength, a test for durability should be developed. This test would eventually replace the strength test in measuring concrete quality.

Variability of Plastic-Concrete Air Content

Air content of portland cement concrete is one of the most important factors in the durability of pavements and bridge decks. Not only is it important that the air content be sufficient to prevent damage from freeze and thaw cycles and low enough to preserve strength, but it is also important that the air be distributed uniformly throughout the mix.

Several States have studied variations in air content and evaluated the performance of different test methods. The data

Table 3.—Historical concrete strength data

	Samples	Mean ($\bar{X}$)	Overall standard deviation (σ_o)	Coefficient of variation (υ)	Testing error (σ_t)
			28-day cylinders—class *A* concrete		
	Number	*p.s.i.*	*p.s.i.*	*Pct.*	*p.s.i.*
Project number:					
1 [1]	536	4,524	396	8.8	175
2	292	4,881	540	11.1	207
3	96	5,686	544	9.6	185
4	192	5,527	566	10.2	155
5	196	5,098	577	11.3	[2] 48
6	112	4,826	608	12.6	196
7	258	5,469	667	12.2	150
8	320	5,244	674	12.9	158
9	232	5,289	711	13.4	192
10	176	4,927	725	14.7	210
11	126	5,067	732	14.4	162
Average [3]	230	5,140	613	13.1	179
Range	224	860	192	5.1	60
			28-day cylinders—*NS* concrete		
Project number:					
1 [1]	50	4,021	[1] 398	9.9	115
2	340	3,555	550	15.5	93
3	240	4,006	580	14.5	134
4	200	3,474	605	17.4	122
5	240	3,781	670	17.7	96
6	196	4,192	729	17.4	160
7	148	4,213	733	17.4	92
8	94	4,239	774	18.3	87
9	108	3,657	774	21.2	116
10	182	3,674	776	21.1	113
11	156	4,110	776	18.9	70
12	224	4,179	825	19.8	127
13	138	3,941	884	22.4	161
Average [3]	178	3,926	698	17.8	114
Range	246	765	334	7.9	91

[1] Values not included in range calculations.

[2] Statistical outlier—not included in calculation of range or average.

[3] Averages are not weighted and include all values except the outlier.

Table 4.—Production source comparison, historical concrete strength data

Source	Concrete class	Samples	Mean strength ($\bar{X}$)	Pooled standard deviation ($\bar{\sigma}$)	Coefficient of variation (υ)
		Number	*p.s.i.*	*p.s.i.*	*Pct.*
1	*A*	480	4,799	593	12.4
	NS	218	4,091	620	15.2
2	*A*	540	4,906	672	13.7
	NS	360	4,210	660	15.7
3	*A*	440	5,054	585	11.6
	NS	308	4,031	584	14.5

shown in table 5, submitted by the State of New York, are representative of this research. The tests proved that, in the State of New York, truck mix concrete was more variable with respect to air than was paver mix or central mix concrete. The tests also showed that the air content measured by the Chace meter was considerably higher than that measured by the pressure meter. As tests with the Chace meter are faster than those with the pressure meter, considerable interest exists in determining the number of Chace meter tests that would give an average that has the same degree of precision as an average based on a lesser number of pressure meter tests.

The equivalency of the Chace and pressure meters can be determined by a comparison of the sampling and testing variance of each in the following manner:

$$\frac{\sigma_{st}^2(C)}{n} = \frac{\sigma_{st}^2(P)}{1}$$

In project 2, table 5, this results in the following equivalency:

$$\frac{.354}{n} = \frac{.166}{1} \therefore n = 2.1$$

or

2 Chace tests = 1 pressure test

Equivalencies computed for the other projects shown in table 5 ranged from one to 20; six of the 10 were below four, indicating that averages based on four Chace meter tests may be suitable for control purposes. However, the wide variation in these results indicate that the actual equivalency of the two tests depends somewhat on the operator's dexterity; and, consequently, it may be necessary to establish operator equivalencies to provide sufficient confidence for the control of air content by the Chace meter in any test.

The data on air content presented in table 6 was reported by West Virginia (*4*). The data indicate that in measuring air content, there is good agreement between the Chace and Roll-A-Meter. Calculation of the equivalency of the two tests indicates that in West Virginia, four Chace tests will adequately substitute for one Roll-A-Meter test.

Variability of Concrete Consistency

The consistency of plastic concrete, as measured by slump cone or Kelly Ball tests, is a measurement of the workability of the mix and an indicator of the water content. However, consistency is no direct measurement of the water content, as air, gradation, and temperature also affect the consistency. The results of these tests therefore are a good indicator of the uniformity of the mix, and relate to a combination of these factors rather than to any one of them.

Results of studies by several States of concrete consistency, as measured by the slump cone, are presented in table 7. The data indicate little difference in the variability in slump among the several methods of concrete production, although there is considerable difference from project to project. The data also indicate that actual material variation contributes more to the overall variation than do sampling and testing.

Data from reports by West Virginia (*4*) and California (*6*) on studies to evaluate the Kelly Ball test are shown in table 8. These data and that from other sources indicate that the

Table 5.—Air content of plastic concrete, research data

Project No.	Mixer type and use	Observations	Test method	Testing variance (σ_t^2)	Sampling variance (σ_s)	Material variance (σ_a^2)	Standard deviation (σ)	Mean ($\bar{X}$)
		Number		*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
1	Central mix	216	Pressure	0.043	0.07	0.80	0.95	4.55
	Paving	216	Chace	0.20	0.545	0.65	1.18	6.39
2	Central mix	200	Pressure	0.126	0.04	0.50	0.82	5.91
	Paving	200	Chace	0.22	0.134	0.70	1.026	7.42
3	E–34 paver	200	Pressure	0.067	0.05	0.48	0.77	5.14
	Paving	200	Chace	0.15	0.10	0.33	0.76	7.38
4	Truck mix	200	Pressure	0.16	0.29	1.74	1.48	7.90
	Structural	200	Chace	0.335	0.15	1.38	1.36	10.2
5	Truck mix	204	Pressure	0.035	0.04	1.27	1.16	6.11
	Structural	204	Chace	0.45	0.99	0.48	1.39	8.75
6	Central mix	200	Pressure	0.06	0.105	0.34	0.71	6.18
	Paving	200	Chace	0.256	0.137	0.40	0.89	7.39
7	E–34 paver	200	Pressure	0.08	0.08	0.39	0.74	4.94
	Paving	200	Chace	0.14	2.55	0.96	1.32	6.41
8	E–34 paver	200	Pressure	0.047	0.04	0.70	0.89	4.82
	Paving	200	Chace	0.26	0.25	1.02	1.24	6.51
9	Truck mix	200	Pressure	0.05	0.135	2.39	1.60	5.80
	Paving	200	Chace	0.43	0.325	1.79	1.60	8.43
10	Truck mix	200	Pressure	0.09	0.136	1.64	1.37	6.07
	Structural	200	Chace	0.28	0.24	1.71	1.49	6.33

Table 6.—Portland cement concrete pavement air content, research data

Project No.	Observations	Test method	Testing variance (σ_t^2)	Sampling variance (σ_s^2)	Material variance (σ_a^2)	Standard deviation (σ)	Mean $(\bar{X})$
	Number		*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
1	176	Roll-A-Meter	0.102	0.133	1.352	1.21	5.8
	141	Chace	0.334	0.233	0.565	1.02	5.06
2	104	Roll-A-Meter	0.109	0.608	0.000	0.83	5.8
	104	Chace	0.470	0.710	0.000	1.00	5.44
3	200	Roll-A-Meter	0.153	0.362	1.042	1.24	5.1
	196	Chace	0.531	0.769	0.913	1.30	5.16
4	172	Roll-A-Meter	0.126	0.248	0.191	0.71	5.1
	154	Chace	0.271	1.026	0.08	1.09	4.85
5	192	Roll-A-Meter	0.143	0.110	1.229	1.16	5.0
	198	Chace	0.249	0.148	1.36	1.32	5.54

Table 7.—Variability of concrete consistency, slump cone method

Project No.	Observations	Testing variance (σ_t^2)	Sampling variance (σ_s^2)	Material variance (σ_a^2)	Overall standard deviation (σ_o)	Mean $(\bar{X})$	Specification limits
State 1							
	Number	*Inch*	*Inch*	*Inch*	*Inch*	*Inches*	*Inches*
[1] 1	184	0.16	0.04	0.26	0.68	2.44	0.5–3.5
[2] 2	200	0.13	0.02	0.45	0.80	1.5	0.5–3.5
[2] 3	300	0.25	0.09	0.46	0.89	2.76	0.5–3.5
State 2							
[2] 1	216	0.074	0.00	0.15	0.47	2.04	....
[4] 2	200	0.06	0.06	0.37	0.70	1.86	....
[3] 3	200	0.08	0.025	0.42	0.73	2.34	....
[5] 4	200	0.027	0.012	0.206	0.495	1.77	....
[5] 5	204	0.066	0.03	0.305	0.633	2.37	....
[6] 6	200	0.033	0.034	0.14	0.456	2.12	....
[7] 7	200	0.084	0.086	0.20	0.609	2.41	....
8	200	0.158	0.047	0.50	0.844	2.26	....

[1] Pavement concrete, truck mix.

[2] Pavement concrete, truck mix, slipform.

[3] Pavement concrete, central mix, screw spreader.

[4] Concrete base, central mix, slide spreader.

[5] Structure concrete, truck mix.

[6] Pavement concrete, central mix, slipform.

[7] Pavement concrete E–34 paver.

Table 8.—Methods of measuring consistency of plastic concrete, research data

Project No.	Samples	Test	Mean $(\bar{X})$	Standard deviation (σ)	Testing variance (σ_t^2)	Sampling variance (σ_s^2)	Material variance (σ_a^2)
West Virginia							
	Number		*Inches*	*Inches*	*Inches*	*Inches*	*Inches*
1	200	Slump	2.4	0.8	0.095	0.095	0.52
		Kelly Ball [1]	2.4	0.7	0.108	0.062	0.30
California							
1	[2] 200	Kelly Ball [3]	3.69	0.91	0.08	0.14	0.61
2	[2] 200	.....do	3.85	0.94	0.22	0.04	0.63
3	[2] 200	.....do	4.00	1.27	0.13	0.10	1.39
4	[4] 200	.....do	1.74	0.65	0.32	0.00	0.10

[1] Converted to inches of slump. Conversion factor—Slump inches = 0.59 Kelly Ball + 1.02.

[2] Structural

[3] Converted to inches of slump by calibration—1-inch penetration of ball indicates 2 inches of slump.

[4] Pavement.

Kelly Ball test is valid for measuring the consistency of concrete when three readings are averaged to obtain one test value as required in the standard method.

Aggregate Size Variation

One factor stressed in concrete specifications and in concrete-production control is gradation of coarse and fine aggregates. Research shows that, within projects and between projects, there is considerable variation in size distribution of aggregates in the mix. In fact, aggregates-size-distribution specifications are seldom complied with. Part 5 will contain a detailed report on the gradation of concrete aggregates, but table 9 is included here to illustrate the variation. The data from project 1 indicate little variation of either material or of sampling and testing; however, the data from projects 2 and 3 indicate a large material variance, and it is probable that the specification limits were exceeded on many individual tests.

Variability in Pavement Thickness

Pavement life expectancy is based on estimated traffic and pavement design thickness. There is still argument as to

Table 9.—Analysis of variance, intermediate aggregate, percent passing ¾-inch sieve

Project No.	Range (R)	Mean ($\bar{X}$)	Material variance (σ_a^2)	Sampling variance (σ_s^2)	Testing variance (σ_t^2)	Standard deviation (σ)	Samples
	Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
1	79–98	92.60	4.25	0.00	8.12	3.52	200
2	33–89	69.09	122.92	5.59	4.54	11.54	200
3	34–92	71.52	124.04	9.31	24.46	12.56	200

Table 10.—Summary of statistical results on thickness of concrete payment

	Samples	Mean ($\bar{X}$)	Overall variance (σ^2)	Standard deviation (σ)	Minimum	Maximum
8-inch uniform thickness						
Project number:	*Number*	*Inches*	*Inch*	*Inch*	*Inches*	*Inches*
1	34	8.66	0.192	0.435	7.63	9.53
2	39	8.42	0.171	0.415	7.61	9.13
3	48	8.35	0.040	0.200	7.86	8.80
4	58	8.36	0.077	0.276	7.76	9.49
5	61	8.05	0.035	0.185	7.66	8.59
6	66	8.11	0.089	0.300	7.46	8.78
7	73	8.06	0.112	0.333	7.58	9.58
Pooled values	....	8.29	0.088	0.300	....	....
9-inch uniform thickness						
Project number:						
1	35	9.25	0.046	0.210	8.93	9.67
2	51	9.19	0.121	0.350	8.55	10.10
3	58	9.28	0.048	0.220	8.84	9.99
4	65	9.18	0.060	0.240	8.78	9.92
5	74	9.20	0.185	0.430	8.69	11.69
6	88	9.11	0.029	0.170	8.85	9.66
Pooled values	....	9.20	0.083	0.290	....	....
10-inch uniform thickness						
Project number:						
1	64	10.38	0.061	0.240	9.41	10.91
2	124	10.34	0.079	0.280	9.82	11.48
3	132	10.35	0.079	0.230	9.75	10.94
4	141	10.28	0.083	0.290	9.63	11.27
Pooled values	....	10.34	0.069	0.270	....	....

whether the design of the pavement should be based on minimum thickness or average thickness; however, as in all structures, stresses are concentrated at the weaker points, and it is axiomatic that large variations in thickness are detrimental to the pavement. Uniformity of thickness will promote better slab action and, therefore, prolong pavement life.

Variation in concrete pavement thickness is shown by the data of figure 2, which is based on a report by the State of Michigan (7). The data depicted represent 656 cores taken from 15 projects from 1959–61. The historical data in table 10, extracted from a report by Louisiana (8), substantiates the variations shown in figure 2.

The variation shown in table 11 is from a statistical study of pavement thickness by the State of Oklahoma (9). The thicknesses were measured by a probe inserted in the plastic concrete. Little variation was exhibited in project 3; but as the mean was below the specified thickness, the pavement life expectancy was less than desired. Project 3 had a probable range of thickness from 7.8 to 10.2 inches, resulting in weak areas that would probably reduce the life of the pavement.

According to the high average concrete thicknesses reported in the Michigan and Louisiana studies, an excess of concrete is being placed by the contractors to avoid penalties. This same high variability-high average thickness relation has also been reported in other studies of thickness. Better control of placement not only could provide savings in concrete, but also produce pavement that is capable of better performance. Moreover, the development of a standard method for measuring the depth of plastic concrete, as placed, would aid in the control of thickness and eliminate expensive coring of the hardened pavement.

Variation in Portland Cement

The production of portland cements is being closely controlled by producers, according to the historical data on chemical analyses reported by several States. It is evident that State highway departments can reduce the testing of portland cement to at least the level recommended by ASTM in section 6 of ASTM-C-183-65T.

Conclusions

Variations in what is generally considered good construction have been shown by the research summarized here. However, the variations are of considerable magnitude and could be important factors in the performance of concrete structures. An awareness of these variations is insufficient; research must be undertaken to evaluate their effect and to develop procedures by which they can be reduced.

In many test results, much of the measured variation could be attributed to sampling and testing methods and procedures, and therefore the *real* variation may not be as large as results indicate. One of the major needs in concrete production is the development of better methods to measure the quality attributes of the concrete and the ingredients incorporated therein. Furthermore a clear delineation of responsibility should result in a more uniform product. It is the contractors' responsibility to produce quality material and the States' responsibility to measure the quality produced. Better measurement performance by the State highway departments will allow a more accurate estimate of product quality and provide a better basis for enforcement of the specification requirements. This approach can result only in a product that is more uniform in character and has improved performance expectancy.

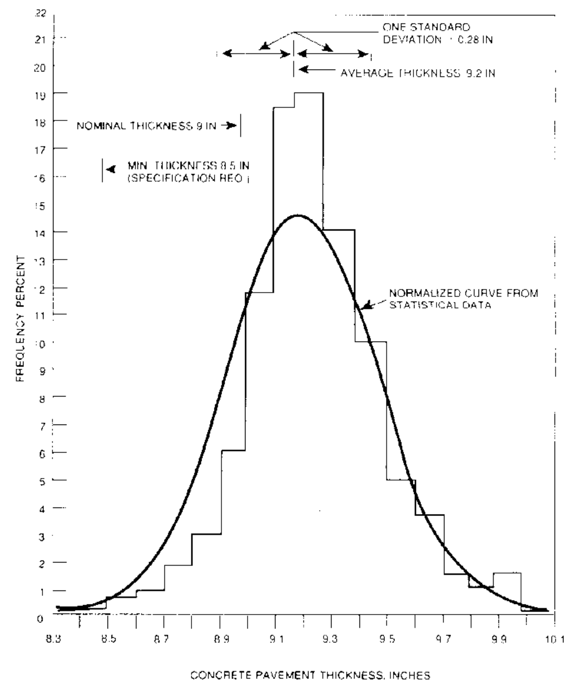


Figure 2.—Frequency distribution of concrete pavement thickness, 1959-61.

Table 11.—Variation in pavement thickness, probe method

Project No.	Observations	Standard deviation (σ)	Mean ($\bar{X}$)	Specification
	Number	*Inch*	*Inches*	*Inches*
1	72	0.3	8.5	8.0
2	95	0.1	8.9	9.0
3	100	0.4	9.0	9.0

REFERENCES

(1) Reflections on Standard Specifications for Structural Design, by Alfred M. Freudenthal, Transactions of the American Society of Civil Engineers, vol. 113, 1948, pp. 269–293.

(2) Designing Specifications—A Challenge, by Edward A. Abdun-Nur, Journal of the Construction Div., Proceedings of the American Society of Civil Engineers, vol. 91, May 1965, pp. 29–44.

(3) Variability of Portland Cement Concrete, by Howard H. Newlon, Proceedings of the National Conference on Statistical Quality Control Methodology in Highway and Air Field Construction, University of Virginia, May 1966, pp. 259–281.

(4) Determination of Statistical Parameters for Highway Construction, Research Project No. 18, The State Road Commission of West Virginia, July 1968.

(5) A Study in the General Field of Quality Control Engineering, The State Road Department of Florida, September 1965.

(6) A Statistical Analysis of the Kelly Ball Test, State of California Division of Highways, Research Report No. M&R 631133–4, October 1966.

(7) Highway Quality Control Program, Statistical Parameters, Michigan Department of Highways, Research Report No. R-572, March 1966.

(8) Quality Control Analysis, Part III, Concrete and Concrete Aggregates, Louisiana Department of Highways, Research Report No. 24, Research Project 63–IG, November 1966.

(9) Statistical Quality Control of Portland Cement Concrete Pavements, Oklahoma Department of Highways, Study No. 64–02–2, June 1968.

Part 4— Variations of Bituminous Construction

**Reported by EDWIN C. GRANLEY,
Highway Research Engineer,
Materials Division**

Introduction

In 1879, a natural asphalt from Trinidad was used on a street project in Washington, D.C., marking the first modern use in this country of a bituminous material in road construction. The practice of treating the uppermost surface of roads with a thin bituminous overlay, such as that used on this street project, continued. The advent of the automobile, bringing with it the production of gasoline and its byproduct, bitumen, brought the material into widespread use, first as a dust preventative and later as a binder for asphaltic concrete.

In the early stages of bituminous mixture development, many of our present specifications and tests were developed to guide contractors and to provide rules for acceptance. Initially, one of the major functions of a specification was to supply technological instructions to the contractor and field engineer. It was necessary to specify exactly how to produce the mixture, how to place it, and how to compact it. Now, the industry has progressed so well that the States soon should be able to specify characteristics of the final product in terms of measurable parameters and to accept it when test results indicate that desired characteristics have been obtained. Before this goal can be reached, however, some problems must be overcome and the ultimate degree to which *end results* specifications can be used in bituminous construction must be determined. Nevertheless, progress is being made in changing from the contractor-State-control construction team to the true contractor-control and State acceptance concept. In shifting responsibilities it is important that acceptance plans protect both the contractor and the State.

To determine the quality characteristics of current construction, many States have been measuring variations in accepted bituminous production. Most of them have followed the guidelines developed by the Bureau of Public Roads Quality Assurance Task Force. The studies patterned after these guidelines not only are providing estimates of the quality of construction, but also are isolating causes of variation.

Data from these studies are revealing the following significant results:

- Variability, indicated by the standard deviation, σ, is itself a variable, and a set value for a standard deviation applicable to the process cannot always be assigned.
- Calculations of the amount of material, or construction component, within present tolerance limits often indicate a considerably lower percentage within the tolerance limits then is expected.
- Test variation, or test error, is often an important factor affecting acceptance or rejection of the material.

Many laboratory-designed tests and sampling plans now being used for on-the-job control and acceptance are inadequate. These devices, developed for 1940 production rates, are still being used to attempt to control and accept bituminous production that exceeds 4,000 tons a day.

Research results obtained in studies of construction variation by several State highway departments have been summarized in this part. Compiling data from bituminous hot-mix projects throughout the country is like putting together a jigsaw puzzle in which the pieces never quite fit and even some are missing. Certain editorial privileges and mathematical manipulations were used to present data uniformly. Sometimes statistical rules were not strictly adhered to. For example, standard deviations, σ, were presented as an arithmetic *average* of individual project results. Components of variance were similarly handled. The term *averages* for data of this type was used to avoid ambiguity with other averaged data. Properly, variance data, in which the square root is directly or indirectly involved, should be *pooled*.

Pooling consists of summing the squared standard deviations (variance) multiplied by the number of test results per project, n, less one $\Sigma\ [(\sigma^2)\ (n - 1)]$, dividing by the total number of test results from all the projects, n, less the number of projects, N, $(\Sigma n - \Sigma N)$, and extracting the square root. The *pooled* standard deviation for the No. 4 sieve in table 1 was 3.56 percent, compared with 3.51 percent obtained from an arithmetical average. From an engineer's standpoint the difference, 3.56 − 3.51 (= 0.05) is considered insignificant. Similar comparisons for other *average* standard deviations, σ, showed a similar difference. This insignificant difference is to be expected, as each project value was obtained from approximately 200 test values and a standard test procedure.

Aggregates

Aggregate represents the largest percentage of any ingredient in a bituminous mixture; consequently, aggregate characteristics significantly control the characteristics of the pavement mixture.

Laboratory research and field experience indicate that, although gradation within the confines of a rather wide grading band is necessary to produce a high-quality product, no single gradation can be adopted as the ideal one for bituminous mixtures. The gradation to be specified depends on the type of surface desired. The maximum size stone used in the pavement is also influenced by the availability of aggregates. The best combination of various sizes then becomes a design problem leading to the establishment of a job-mix formula. The job-mix formula also includes the desired asphalt content.

Under present practice, the State often accepts responsibility for determining the job-mix formula. Once the job-mix formula is established and approved by the engineer, it becomes the targeet or central value for process control. A tolerance is usually included to account for normal variability of materials or processes.

Variations in aggregate gradation

According to the research studies being conducted, randomly selected samples, taken independently of control samples, usually show deviations in gradations that often are larger than the specification tolerances. Summaries for each aggregate gradation in presently accepted construction and

Table 1.—Average of aggregate gradation data from extraction tests

Sieve size	Average standard deviation ($\bar{\sigma}$) of percent passing	Shift of average ($\bar{X}$) from job mix target	Average variance components as a percent of total variance (σ_o^2)			Computed average compliance with job mix tolerances
			Testing	Sampling	Material	
Surface mixes, 22 projects						
	Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
¾ in or ½ in	1.43	1.70	72	4	24	99
⅜ in	2.49	1.73	29	31	40	93
No. 4	3.51	2.95	12	18	70	78
No. 8 or 10	2.81	2.45	10	15	75	77
No. 20 or 30	1.74	2.10	13	18	69	87
No. 40 or 50	1.37	1.72	18	15	67	87
No. 80 or 100	1.00	1.44	17	11	72	82
No. 200	0.94	1.43	21	14	65	74
Average .... pct	1.91	1.94	24	16	60	85
Base or binder mixes, 6 projects						
¾ in or ½ in	4.33	1.66	65	13	22	83
⅜ in	4.93	5.88	55	30	15	60
No. 4	3.92	2.03	46	17	37	76
No. 8 or 10	2.53	1.81	19	13	68	50
No. 20 or 30	2.17	2.22	25	28	47	81
No. 40 or 50	1.67	1.63	23	31	46	84
No. 80 or 100	1.15	1.23	30	30	40	97
No. 200	0.88	1.02	21	14	65	74
Average .... pct	2.70	2.19	36	21	43	76

their relations to specified tolerances are shown in tables 1 and 2 and in figures 1, 2, and 3. Data for surface course mixes are included from 22 projects in eight States and for binder or base mixes from six projects in five States.

A consolidation of gradation data for aggregates from extraction test results is shown in table 1. These data were obtained on samples taken independently of those used for job control and acceptance. Departure of averages, $\bar{X}$, from

Table 2.—Averages of surface course aggregate data from extraction tests on 22 projects

Sieve size	Average standard deviation ($\bar{\sigma}$) of percent passing			Deviation of average ($\bar{X}$) from job mix target			Suggested AASHO tolerance limits	Computed compliance with suggested AASHO tolerance limits		
	⅓ of jobs having least variable σ	All jobs	⅓ of jobs having most variable σ	⅓ of jobs having least variable σ	All jobs	⅓ of jobs having most variable σ		⅓ of jobs having least variable σ	All jobs	⅓ of jobs having most variable σ
	Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
¾ in or ½ in	0.81	1.43	2.32	0.53	1.70	2.62	±7	100	100	95
⅜ in	1.71	2.49	3.46	1.06	1.73	2.36	±7	100	98	91
No. 4	2.33	3.51	4.52	2.22	2.95	4.83	±7	98	87	68
No. 8 or 10	1.90	2.81	3.85	1.68	2.45	4.04	±4	89	71	49
No. 20 or 30	1.32	1.74	2.24	1.70	2.10	3.06	±4	96	86	66
No. 40 or 50	0.93	1.37	1.82	1.41	1.72	1.23	±4	100	95	94
No. 80 or 100	0.65	1.00	1.36	0.65	1.44	1.48	±4	100	99	97
No. 200	0.51	0.94	1.45	1.43	1.43	1.74	±2	87	73	57
Average .. pct .	1.29	1.93	2.64	1.33	1.94	2.67	....	96	89	77

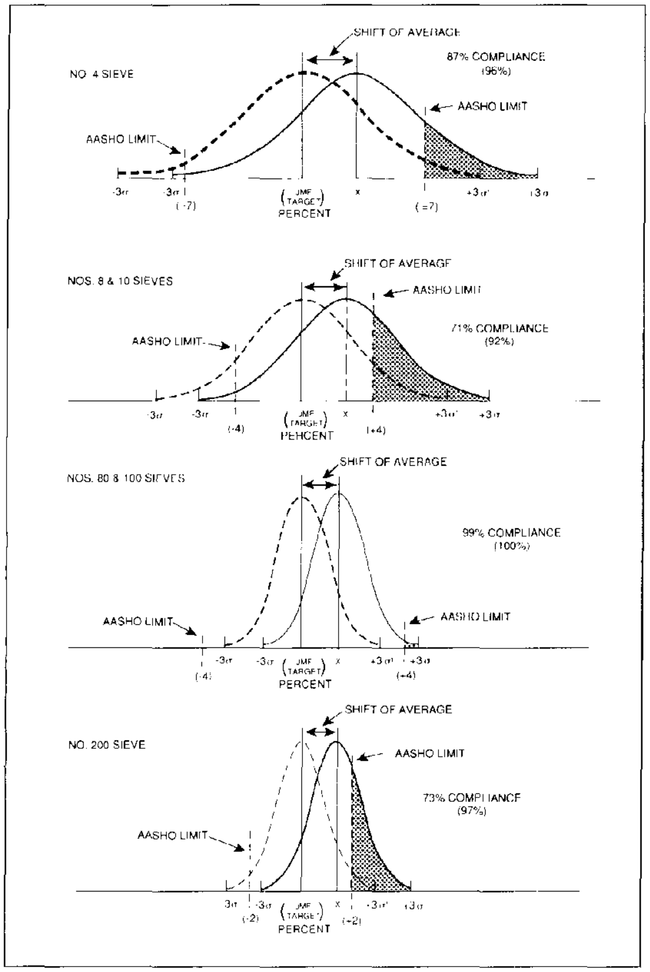


Figure 1.—Computed 3σ limits on target values of job mix formula for aggregate gradation—surface and binder or base courses.

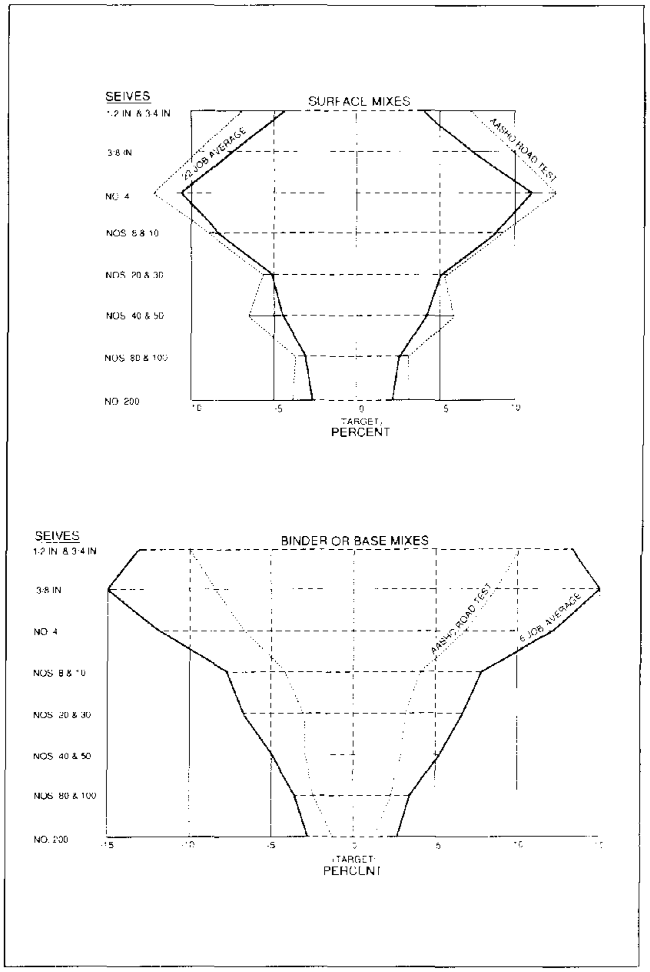


Figure 2.—Theoretical-frequency-distribution curves of gradation data for selected sieves.

the job-mix formula for individual jobs were about evenly divided below and above this target.

An analysis was conducted to determine the components of variance that could be attributed to sampling, σ_s^2, testing, σ_t^2, and materials, σ_a^2. For surface material, combined testing and sampling variances ($\sigma_t^2 + \sigma_s^2$) are shown to be in the range of 25–35 percent of the total variance, (σ_o^2), for sieve No. 4 and smaller sieves. For the larger sizes of either surface or base course materials, the combined sampling and testing error was a significantly larger proportion of the total variance. For the base course materials, even smaller sizes showed large sampling and testing variances. The statistically computed average percent compliance to States' job-mix formula and tolerances are also shown.

Further analysis of the data from the construction projects for surface course materials is given in table 2. These data provide a summary of variations from the least variable one-third, and the most variable one-third of the projects, as well as the average for the total. Also, statistically computed percent compliances with suggested tolerance limits of the AASHO Guide are shown instead of computed conformance to job-mix tolerances. In general the most variable projects show average standard deviations, $\bar{\sigma}$, of about twice the corresponding values for the least variable projects.

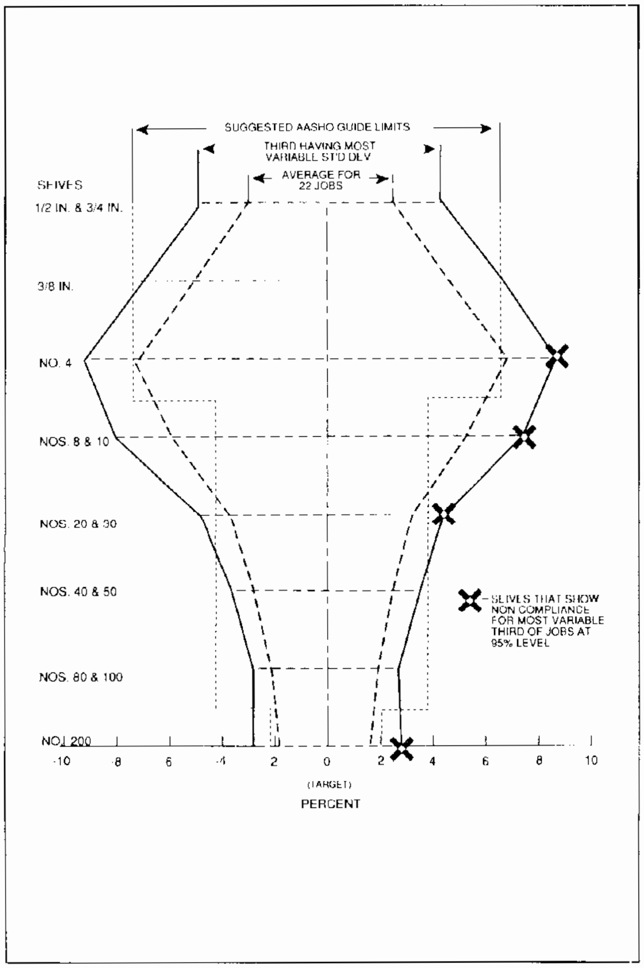


Figure 3.—Comparison of surface-course-gradation 2σ limits and suggested AASHO guide limits.

The plus or minus three average standard deviations, $\pm 3\bar{\sigma}$ (table 1) for both surface and binder or base courses are plotted in figure 1. The bulged shape, or the largest spread, emerges at the No. 4 sieve for surface courses, and at the ⅜-inch and larger sieves for binder and base courses. Average variations for base and binder courses are about ⅓ larger than those for surface courses. Superimposed on each diagram are $\pm 3\sigma$ values for each type of mix from the AASHO Road Test (*1*).[1] Because the construction of the AASHO Test Road was very carefully controlled, these data are considered a solid base with which to compare research data. For average surface course data, almost perfect agreement is shown with the AASHO Road Test results. There is no apparent reason why AASHO Road Test gradation data show less variation (smaller standard deviation) for base courses than for surface courses.

Both the standard deviation and the shift of the average, $\bar{X}$, from the job-mix target affect conformance to specifications. The effect of $\bar{X}$ shift from the target value is shown in figure 2, in which values from table 1 were used to compute theoretical normal frequency distributions of four selected sieve size groups. Darkened areas of the tips of the curves represent noncompliance with suggested (*2*) AASHO Guide Specifications. The dashed line curves, which are the same distributions shown by the solid curves, are superimposed on the job-mix target values. The percent compliance with AASHO limits of the superimposed curves is usually much larger than that of the solid curves, as is indicated by the figures in brackets.

Many consider that a reasonable conformance to a specification is met if tests indicate that 95 percent of the material is within the stated tolerance. For a normal distribution this 95-percent conformance level approximates the two standard deviation limits. For the 22 surface course projects, the relation of the spread of gradation represented by $\pm 2\sigma$ from the average to the suggested limits of the AASHO Guide Specification (*2*) is shown in figure 3. The $\pm 2\sigma$ range is also shown for the ⅓ of the jobs with the largest standard deviation. Except for No. 8 and No. 10 sieve group, the $\pm 2\sigma$ range of all jobs are within the AASHO Guide Limit. The $\pm 2\sigma$ ranges for the third with the largest standard deviations are relatively close to the AASHO limits; the most significant deviation occurred again for the No. 8 and No. 10 group. There were also significant deviations for the No. 4 and the No. 200 sieves.

---

[1] Italic numbers in parentheses identify the references listed on p. 44.

The relation of the standard deviation to the average percent retained on each sieve is shown in table 3 for the different groups of surface and base or binder mixes. It will be noted that the standard deviation for surface course mixes seems to be related to the amount of material retained on each sieve, as shown in figure 4.

Asphalt Content

The quality and quantity of asphalt in a pavement mixture largely determines the useful life of the pavement, provided that the pavement has been properly compacted. Too much asphalt in a mixture can cause flushing and rutting of the pavement, and too little asphalt can cause cracking or raveling. Thus, a close control of asphalt content is desirable.

Asphalt content data from extraction tests on 26 surface course mix and seven base course or binder mix projects are arranged in table 4 according to size of standard deviation. Also, the surface course mix projects are grouped by thirds to delineate those with the least variable standard deviation, the middle third, and those with the most variable standard deviation. Shown in separate columns are the plus or minus shift of the job average from job-mix target and statistically computed compliance with ± 0.4, ± 0.6, and ± 0.8 percent tolerances, respectively.

The average $\bar{\sigma}$ of extracted asphalt for surface mix projects was 0.28 percent. The average for binder or base mix projects was 0.35 percent. The computed $\pm 3\sigma$ limits for 20 of the 26 surface course projects, in which the job-mix target

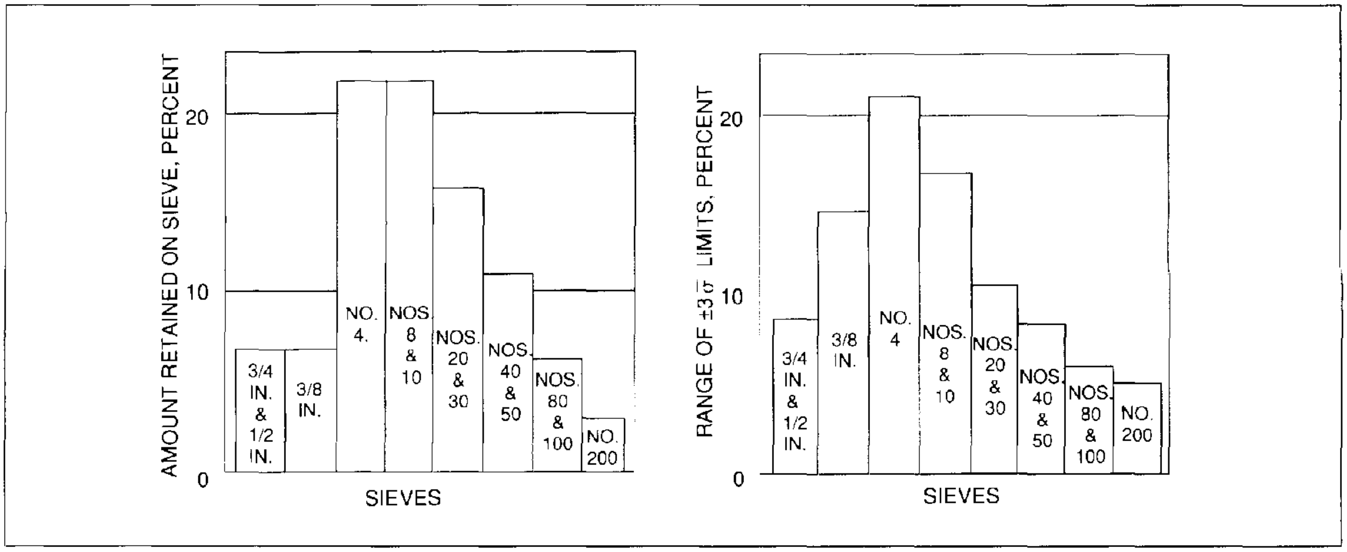


Figure 4.—Comparison of average amount retained on sieves and range of $3\bar{\sigma}$ limits—surface course projects.

Table 3.—Average sieve data from extraction test

Sieve size	Average amount passing [1]	Average amount retained on indicated sieve size [1]	Range of ±3 average standard deviation limits				Suggested AASHO guide (±) limits
			Surface course mixes			Base or binder mixes all jobs	
			Least variable third	All jobs	Most variable third		
	Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
¾ in or ½ in	93	7	[2] 2.4	[2] 4.3	[2] 7.0	13.0	7
⅜ in	86	7	[2] 5.1	7.5	10.4	14.8	7
No. 4	64	22	[2] 7.0	10.5	13.5	11.8	7
No. 8 or 10	42	22	5.7	8.4	11.5	7.6	4
No. 20 or 30	26	16	[2] 4.0	5.2	6.7	6.5	4
No. 40 or 50	15	11	[2] 2.8	4.1	5.5	5.0	4
No. 80 or 100	9	6	[2] 1.9	[2] 3.0	4.1	[2] 3.5	4
No. 200	6	3	[2] 1.5	2.8	4.4	2.6	2

[1] For surface mixture only.

[2] Within AASHO guide recommended tolerance limits.

Table 4.—Bituminous content data from extraction tests

Job No.	Standard deviation (σ)	Shift of average from job mix target		Computed compliance with tolerances from job mix target		
		Below	Above	Suggested AASHO guide tolerance ± 0.4%	Assumed tolerances	
					± 0.6%	± 0.8%
Surface course mixes						
	Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
Least variable third of jobs:						
1	0.12	0.00	0.00	100	100	100
2	0.14	0.31	....	74	98	100
3	0.14	0.04	....	99	100	100
4	0.17	....	0.07	97	100	100
5	0.18	0.22	....	88	98	100
6	0.19	0.44	....	42	80	97
7	0.19	0.13	....	92	99	100
8	0.19	0.20	....	85	98	100
9	0.21	....	0.20	83	97	100
Average, least variable third	0.17	0.18		84	97	100
Middle third of jobs:						
10	0.22	....	....	....	....	....
11	0.22	....	....	....	....	....
12	0.23	0.06	....	93	99	100
13	0.26	....	....	....	....	....
14	0.26	0.15	....	83	96	100
15	0.27	0.22	....	74	92	98
16	0.27	0.30	....	64	87	97
17	0.27	....	....	....	....	....
Average, middle third	0.25	0.18		79	94	99
Most variable third of jobs:						
18	0.33	....	....	....	....	....
19	0.34	....	0.20	68	88	96
20	0.37	0.60	....	29	50	79
21	0.38	....	10.30	51	78	91
22	0.38	....	10.25	58	82	93
23	0.47	....	....	....	....	....
24	0.47	0.35	....	50	68	83
25	0.49	0.23	....	54	73	86
26	0.53	....	0.20	52	71	84
Average, most variable third	0.42	0.30		52	73	87
Average, surface course mixes	0.28	0.22		72	88	95
Base or binder course mixes						
1	0.22	....	....	....	....	....
2	0.27	....	0.33	60	88	96
3	0.28	....	....	....	....	....
4	0.38	....	....	....	....	....
5	0.38	0.24	....	62	81	93
6	0.43	....	0.13	63	82	93
7	0.50	....	....	....	....	....
Average, binder or base mixes	0.35	0.23		62	84	94

value was reported, are shown in descending order in figure 5. Also shown in figure 5 is the shift of the average from the target value. The asphalt content for Project No. 1 was on the target; it was also the only project to show variations that were less than the suggested ± 0.4 percent limits of the AASHO Guide. The three standard deviation limits for individual projects ranged from 0.36 to 1.59 percent. The computed 3 σ limits for AASHO (*1*) and WASHO (*3*) road tests were 0.54 percent and 1.20 percent, respectively. On about ⅔ of the jobs, the job averages were lower than the target (table 4). Only three surface mix jobs complied 100 percent with assumed tolerances from the job-mix formula of ± 0.6 percent, although half the total showed more than 95 percent compliance. Increasing the tolerance to ± 0.8 percent did not appreciably increase the number of jobs having more than 95 percent conformance.

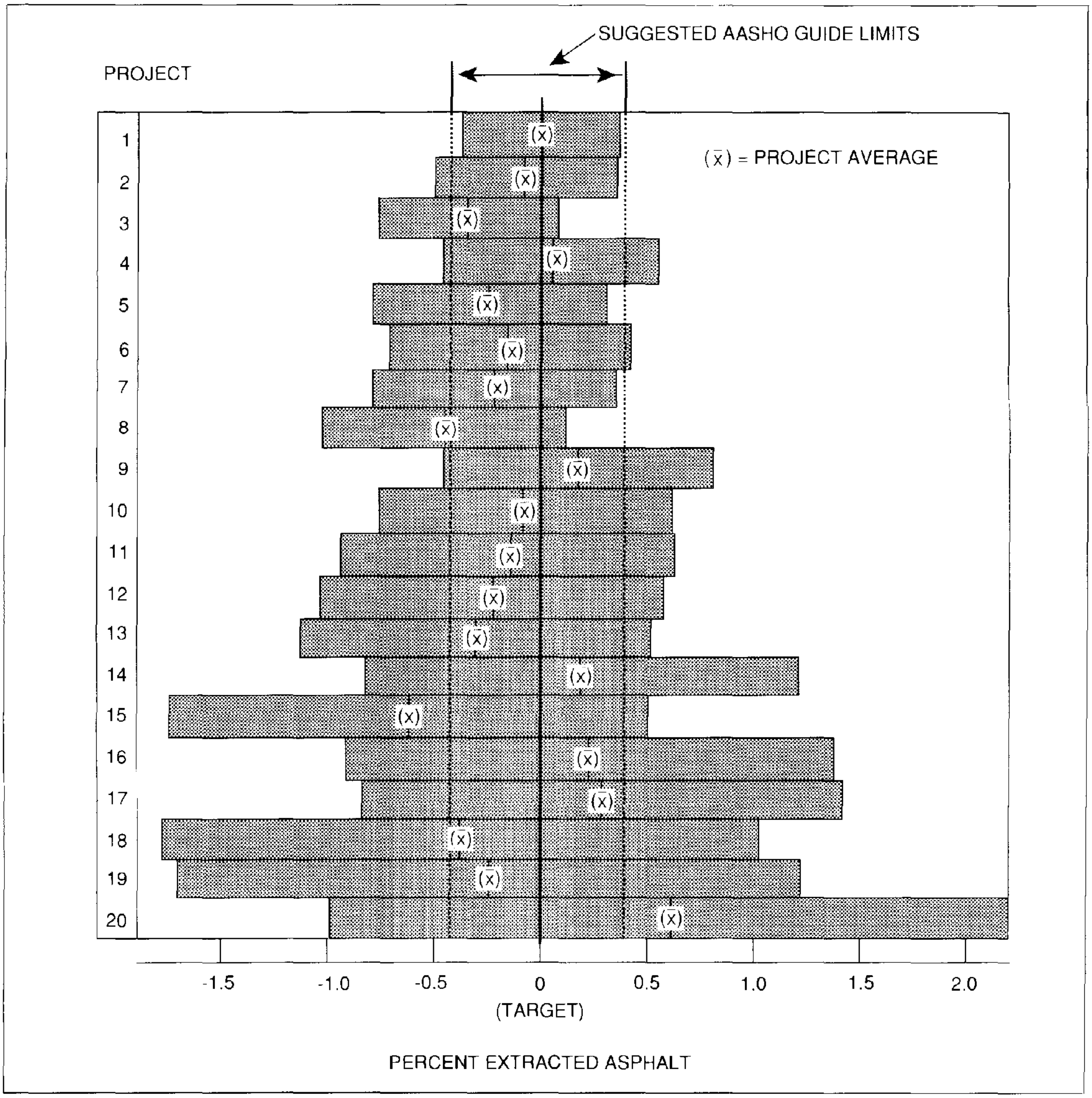


Figure 5.—Computed 3σ limits and shift of average from job-mix formula target for extraction test data of asphalt content—20 surface course projects.

Data for the surface course projects from table 4 are shown in figure 6, grouped into three sections according to standard deviation of asphalt content. The $\bar{\sigma}$ for each group was used to construct the three normal curves, which show that the most variable projects also had the largest shift of the $\bar{X}$ from the target value. This shift indicates a lack of job control that adversely affects both the average and the variance. The computed conformance percentages are based on the AASHO's suggested ±0.4 percent tolerance. In parentheses, beneath the percent conformance for each group, is the computed percentage that would have been obtained if all projects' averages had been on the target value.

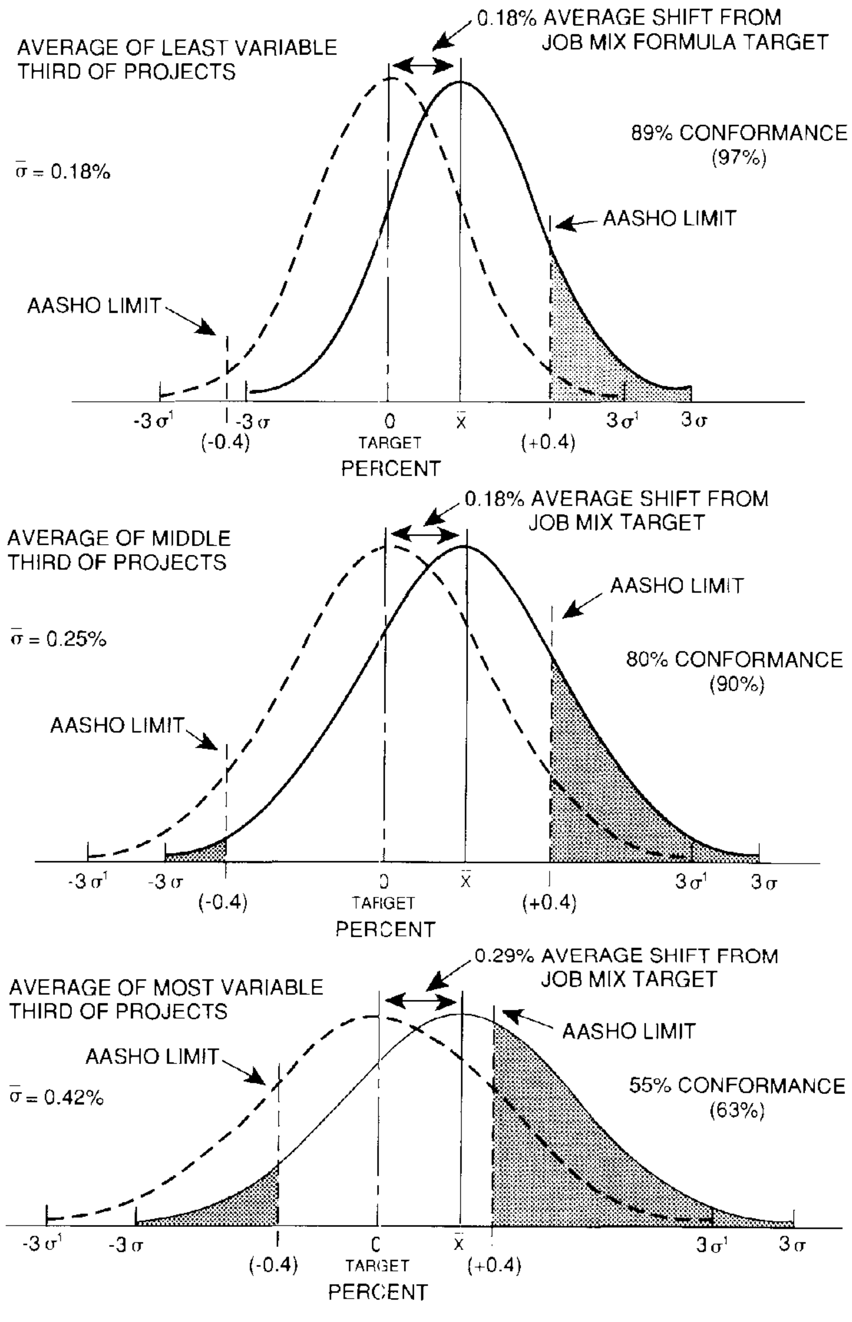


Figure 6.—Normal distribution curves, conformance to suggested AASHO guide tolerance of ±0.4% for asphalt content—three groups of surface course projects.

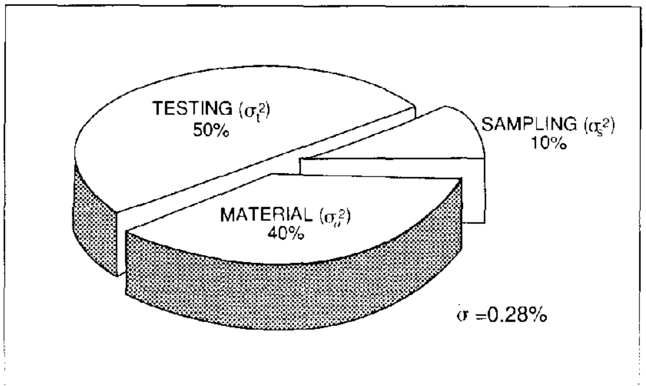


Figure 7.—Average percent of total variance, σ_o^2, attributable to testing, sampling, and material variances for asphalt content extraction tests—23 surface course projects.

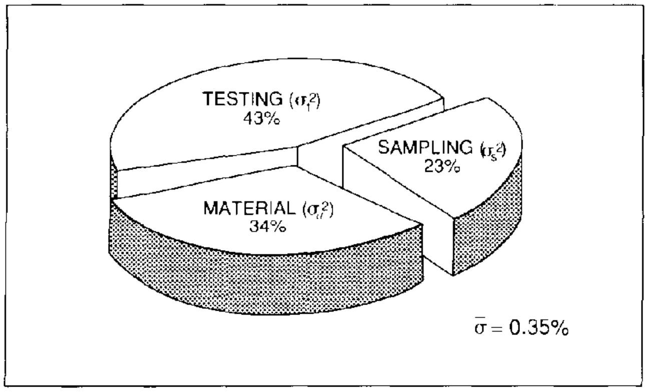


Figure 8.—Average percent of total variance, σ_o^2, attributable to testing, sampling, and material variances for asphalt content extraction tests—6 base or binder course projects.

Testing and sampling variance

Testing, sampling and material variances of asphalt contents for both surface and binder or base-course mixes are shown in figures 7 and 8, respectively. These variations, imply that results of a single extraction test are not a reliable measure of asphalt content. However, the precision of the measurment can be improved by using the average of several results as the test value. Better precision can also be obtained by improved sampling and testing procedures.

Testing Variations

Effect of sampling point

Engineers disagree as to whether the location at which a sample is taken affects test results. According to present practice, extraction test samples usually are obtained from the truck at the plant so that results can quickly be made available. Research has been performed to evaluate the effect of the sampling location. Average test results of samples from the truck and those of core samples from the pavement are listed in table 5. These data from 10 jobs indicated no significant differences between core samples and truck samples. The bar graphs in figure 9 substantiate that the point of sampling does not significantly affect the variances for asphalt content.

Ash correction

The extraction test for determining asphalt content includes an ash correction for insoluble material that passes through the filter. Because field laboratories do not always operate under optimum conditions, it is thought by some that the State should dispense with running the ash correction in

Table 5.—Average sieve data, aggregate residue and asphalt content—from extraction tests of samples obtained from same mix at two locations on 10 projects in three States

	Average deviation ($\bar{\sigma}$)		Average shift of average ($\bar{X}$) from job mix formula target		Average variance components as a percent of total variance (σ_o^2)						Average percent compliance with job mix formula tolerances	
					Testing		Sampling		Material			
Sample location	Truck	Core	Truck	Core	Truck	Core	Truck	Core	Truck	Core	Truck	Core
	Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
Sieve:												
¾ in or ½ in.	1.33	1.69	1.58	1.11	74	32	1	2	25	66	99	100
⅜ in.	2.34	2.42	1.16	0.87	37	39	22	14	41	47	98	98
No. 4	2.89	2.96	1.68	2.14	26	27	21	28	53	45	85	83
No. 8 or 10	2.53	2.58	1.81	2.50	19	21	13	24	68	55	84	87
No. 20 or 30	1.52	1.73	1.59	2.06	12	13	18	17	70	70	92	86
No. 40 or 50	1.45	1.66	1.80	2.00	22	16	6	8	72	76	84	79
No. 80 or 100	1.06	1.09	1.34	1.63	27	21	10	9	63	70	79	74
No. 200	0.98	0.97	1.05	1.26	27	24	11	10	62	66	74	70
Average	1.76	1.88	1.50	1.69	31	24	13	14	56	62	87	85
Asphalt	0.22	0.22	[1] 0.23	[1] 0.22	32	40	11	22	57	38	[1] 61	[1] 63

[1] From six jobs in one state only.

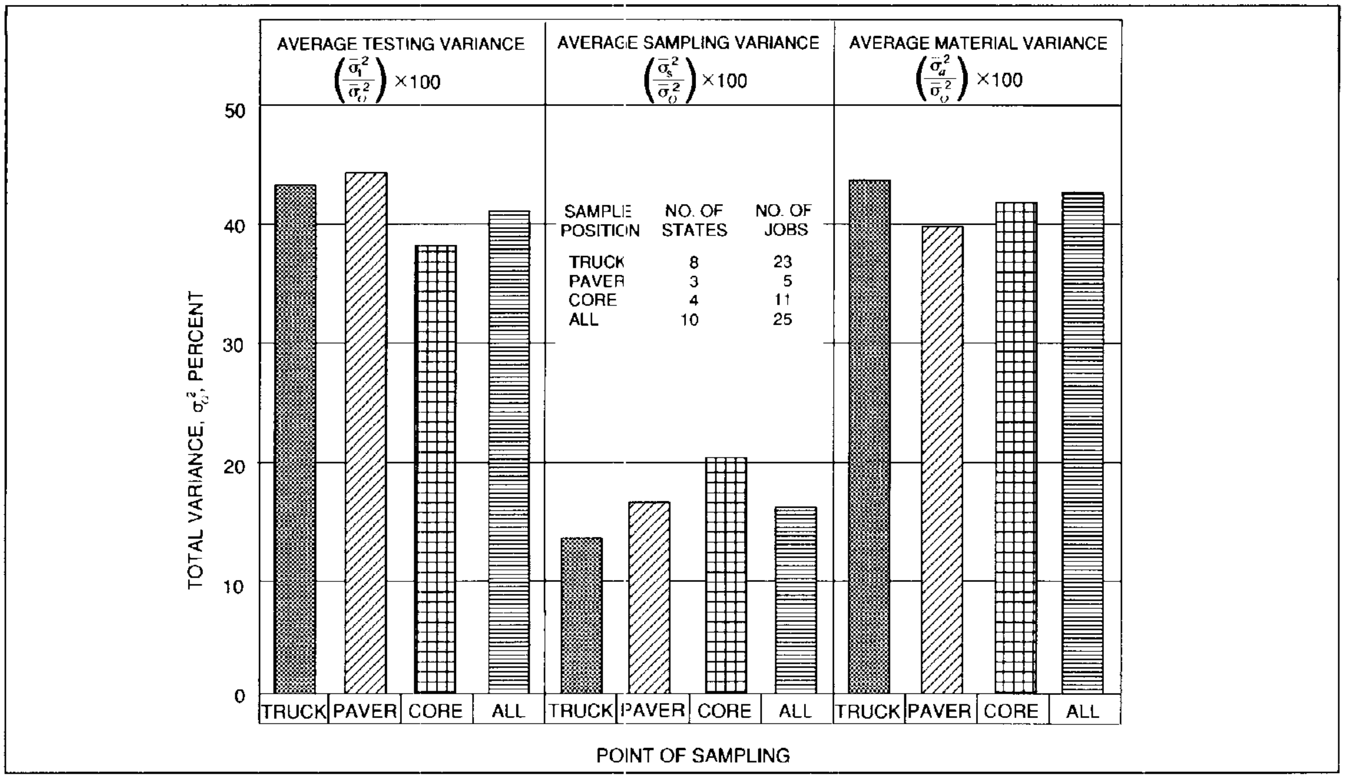


Figure 9.—Average percent of total testing variance, σ_o^2, attributable to testing, sampling, and material variances for asphalt content extraction tests—alternate sampling locations.

field laboratories and substitute constant corrections determined by a central laboratory.

Several studies were conducted to determine ash correction variations in the field. In a Florida report (*4*), field laboratories tests, when compared with central laboratory tests, were shown to be inconsistent. All field laboratories weighed their ash correction residue to the nearest 0.1 gram, instead of to 0.01 gram, apparently because of the sensitivity of available scales. Some corrections were made on the basis of a constant factor per 100 cc. of solvent used in the test. Field laboratories also used more solvent, and the quantities of solvent varied more from test to test than those of the central laboratory.

Central laboratory and field laboratory ash corrections were compared by testing split samples taken from surface and binder mixes on 10 jobs. The results showed that the field laboratories had a smaller $\bar{X}$ ash correction and were, on the average, less variable. On individual jobs, this trend was not so pronounced, as shown by the following tabulation.

Ash correction:

	Grams
Surface mix, average ($\bar{X}$):	
Central laboratory	4.60
Field laboratory	4.10
Binder mix, average ($\bar{X}$):	
Central laboratory	4.95
Field laboratory	4.10
Surface mix, standard deviation (σ):	
Central laboratory	2.48
Field laboratory	1.64
Binder mix, standard deviation (σ):	
Central laboratory	2.28
Field laboratory	1.84

The lower values and less variability of the field tests do not necessarily indicate that the results are more accurate. In ash correction there is always a danger of not obtaining a true aliquot because of ash settlement in the container, which could cause indicated trends.

The variances obtained using both a constant ash correction factor and actual field correction factors were compared for binder and surface mixes using the chi square statistic, χ^2. Neither calculated value of χ^2 reached the critical 5-percent significance level. Statistically, from the Florida report (*4*), "it has not been demonstrated that any significant difference exists." This means both methods will produce the same results. The Florida report further states:

"At present the evidence seems to indicate that if the operation of the extraction test could be improved (specifically: uniformity in devices, amount of solvent used, number of washes employed, speed of rotation, etc.), there is a very good possibility that the running of the ash correction as a field test could be dispensed with and a system devised using a factor assigned by the central laboratory, which would give statistically as good, if not better results, than are being obtained under the present system. Periodic spot checks and inspections of equipment, procedures, etc., would undoubtedly have to be made to ensure that continued high standards of operation were continuously being obtained."

Effect of extraction test equipment, operators, and laboratories

Extraction tests on the 33 projects shown in table 4 were made with either Reflux or Rotorex test equipment. Except for those of two States, all extraction and sieve tests were performed at district or central laboratories. In New Jersey (5), half the extraction tests were made at the central laboratory, and available plant testing equipment and plant inspectors were used to test the remaining half to determine whether any significant testing variability or *variability of testing variability* existed. According to the data in table 6, which is from a report by Afferton (5), testing variance, σ_r^2, for determining asphalt content was more than 15 times greater in field laboratories than in the central laboratory. The statistical test for differences of σ_r^2, using the F ratio at the 5-percent level, showed a high significance of testing variance for both courses.

A comparison test on split samples using both Reflux and Rotorex test equipment was reported in a West Virginia study (*6*). On the basis of t and F statistical tests, no significant differences in standard-deviation variability could be attributed to the type of test equipment. However, in companion studies, in which two sets of samples with known asphalt quantities, two operators, and both sets of equipment were used, it was shown that operator proficiency significantly affected the accuracy of the test results, possibly enough to nullify the smaller standard deviation expected of the Reflux apparatus. In another experiment, in which six operators each used Rotorex equipment to test two samples with known asphalt content (unknown to operators), the operators retained their same numerical order of proficiency.

A Florida study (*4*) also statistically compared field asphalt content determinations of binder and surface mixes made by regular plant inspectors on 10 jobs with central laboratory test results of duplicate samples and by percent of total variance from regression lines for each. Essentially the

Table 6.—Tests for significant variance difference between field and laboratory testing, 5-percent level

Test property	Testing variance (σ_r^2)		Largest variance	F ratio		Is difference significant?
	Laboratory	Field		Computed	Critical	
Top						
	Pct.	*Pct.*		*Pct.*	*Pct.*	
Asphalt content	0.0088	0.1734	Field	[1] 19.70	1.75	Yes.
Stone content	1.5500	2.9040	Field	1.87	1.75	Yes.
Sieve analysis:						
Passing 1 in., retained on ½ in	0.7200	1.0358	Field	1.44	1.75	No.
Passing ½ in., retained on ¼ in	2.1600	6.2827	Field	2.41	1.75	Yes.
Passing ¼ in., retained on No. 10	0.9200	0.7246	Laboratory	1.27	1.75	No.
Passing No. 10, retained on No. 30	0.9000	0.3591	Laboratory	2.51	1.75	Yes.
Passing No. 30, retained on No. 50	1.6400	1.8232	Field	1.11	1.75	No.
Passing No. 50, retained on No. 80	1.1700	0.9429	Laboratory	1.24	1.75	No.
Passing No. 80, retained on No. 200	0.7600	3.0043	Field	1.95	1.75	Yes.
Passing No. 200	0.2900	0.5121	Field	1.76	1.75	Yes.
Bottom						
Asphalt content	0.0111	0.1658	Field	[1] 14.94	1.75	Yes.
Stone content	2.9700	7.9839	Field	2.69	1.75	Yes.
Sieve analysis:						
Passing 1½ in., retained on 1 in	13.8400	19.7247	Field	1.42	1.75	No.
Passing 1 in., retained on ½ in	17.1100	23.1702	Field	1.35	1.75	No.
Passing ½ in., retained on ¼ in	8.0100	9.3947	Field	1.17	1.75	No.
Passing ¼ in., retained on No. 10	3.0000	1.4930	Laboratory	2.01	1.75	Yes.
Passing No. 10, retained on No. 30	0.6200	0.3234	Laboratory	1.92	1.75	Yes.
Passing No. 30, retained on No. 50	0.5400	0.8708	Field	1.61	1.75	No.
Passing No. 50, retained on No. 80	0.4600	0.4770	Field	1.04	1.75	No.
Passing No. 80, retained on No. 200	0.8700	0.9584	Field	1.10	1.75	No.
Passing No. 200	0.3000	0.2867	Laboratory	1.05	1.75	No.

[1] Highly significant at 5-percent level.

same operators were rated good-to-fair and poor-to-very poor in both types of tests, indicating that operator training and constant surveillance is necessary to achieve precise extraction test results.

Control by hot bin sieving

From 1 to 2 hours are required to complete field extraction tests now being used to determine whether bituminous hot mix conforms to the requirements of the job-mix formula. For this reason several State highway departments have been seeking quicker means to ascertain conformance so that remedial action can be taken quickly.

In a New York study (7), it was determined from research comparisons (see table 7) on dry hot bin and extraction sieve tests that dry sieving was more uniform for ½-in., ¼-in., and No. 200 sieve sizes. The extraction test yielded more consistent results for the ⅛-in. thru No. 80 sieves. As accurate printed weights of material used in each batch from each bin were obtained, it was decided to use the more rapid hot bin sieving to control the uniformity of the mix. This test was to be supplemented with a daily extraction test for aggregate passing the No. 80 and No. 200 sieves. According to the hot bin data from the 29 plants in which the tests were performed, anytime that the primary size in the coarse bin—material passing 1-in. sieve and retained on ½-in. sieve in No. 1 bin, and passing ¼-in. sieve and retained on ⅛-in. sieve in 1A bin—fell below 70 percent, the mix generally became nonuniform. By trial it was determined that a 12-percent fluctuation in this quantity from the last test was a practical limit to use in order to avoid exceeding the job-mix formula limits. On the fine aggregate bin, the same tolerance limit was applied to material retained on the No. 20 sieve, because usually about one-half of the fine aggregate was retained on this sieve.

Because of the relation of primary size to the overall conformance to the job formula, the New York State highway department is using this correlation as an indicator of uniformity. The uniformity control test is supplemented by complete hot bin analysis, usually after every fourth test. Thus, one State has been able to shift dependence on extraction test results to a secondary role.

Their inspection manual states:

"In general, production is accepted by obtaining gradation tests results within the limits of a job mix formula. Hot bin analyses and uniformity tests determine the gradation of material larger than the No. 80 sieve. The extraction test is used to determine gradation of material smaller than the No. 80 sieve and also indicates the approximate bitumen content. Actual bitumen content is determined by verifying batch quantities."

Density

Permanence of bituminous pavement depends largely on the degree of compaction obtained. The compaction value is usually expressed as a percentage of either theoretical (voidless) or Marshall density, determined by laboratory tests. Density data from several construction jobs are shown in table 8. It is quite possible that much below-specification density can be attributed to improper rolling patterns. Figure 10 was taken from a report by Kilpatrick and McQuate (8) who reported the following conclusions regarding effect of rolling pattern on density:

"Normal rolling procedures used by roller operators result in wide lateral variations in compactive effort. The number of roller passes applied in the center of the lane is usually from three to six times greater than at the lane edges.

"The lateral pattern of density is similar to the lateral pattern of compactive effort; i.e., high-in-the-middle and low-at-the-edges."

In figure 10, the density pattern across the lane approaches the shape of a normal curve. In a random selection of sample locations across the lane, sites at any distance from the edge have an equal chance of being selected.

Table 7.—Comparison of dry hot bin and extraction results[1]

Sieve size	Total percent passing			Standard deviation (σ)	
	Average hot bin	Average extraction	Average difference[2]	Pooled hot bin	Pooled extraction
	Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*
½ in	99.6	99.6	0.0	0.3	0.5
¼ in	78.6	77.8	0.8	2.5	3.7
⅛ in	47.5	46.2	1.3	3.0	2.3
No. 20	21.0	21.4	–0.4	3.7	2.3
No. 40	13.3	14.7	–1.4	3.3	2.3
No. 80	6.3	7.8	–1.5	1.9	1.3
No. 200	2.8	4.5	–1.7	1.0	1.0
Percent asphalt content	....	6.3	....	....	0.3

[1] Data based on 491 combined hot bin analyses and 491 extraction tests from 29 mix plants during 1962, 1963, and 1964.

[2] Difference is significant at 99 percent confidence level for all sieves except for No. 20, which is significant at 95 percent confidence level.

Table 8.—Average bituminous hot mix density data from research jobs

	Jobs	States	Average standard deviation ($\bar{\sigma}$)		Average ($\bar{X}$)		Average variance components as percent of total variance (σ_o^2)			Percent compliance with State specification
			Core	Loose sample	Core	Loose sample	Testing	Sampling	Material	
Percent of theoretical density (voidless):	*Number*	*Number*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
Surface	15	6	1.57	....	93.1	....	5	19	77	78
Binder	3	2	2.90	....	94.2	....	33	16	51	88
Percent of Marshall density: Surface	12	5	1.53	....	96.0	....	....	....	....	....
Marshall density percent of theoretical density: Surface	10	2	....	0.89	....	96.2	20	12	68	....
Theoretical density (voidless):			*grams/cc.*	*grams/cc.*	*grams/cc.*	*grams/cc.*				....
Surface	10	3	0.013	0.011	2.43	2.46	18	12	70	....
Binder	1	1	0.029	0.013	2.48	2.48	....	....	....	....

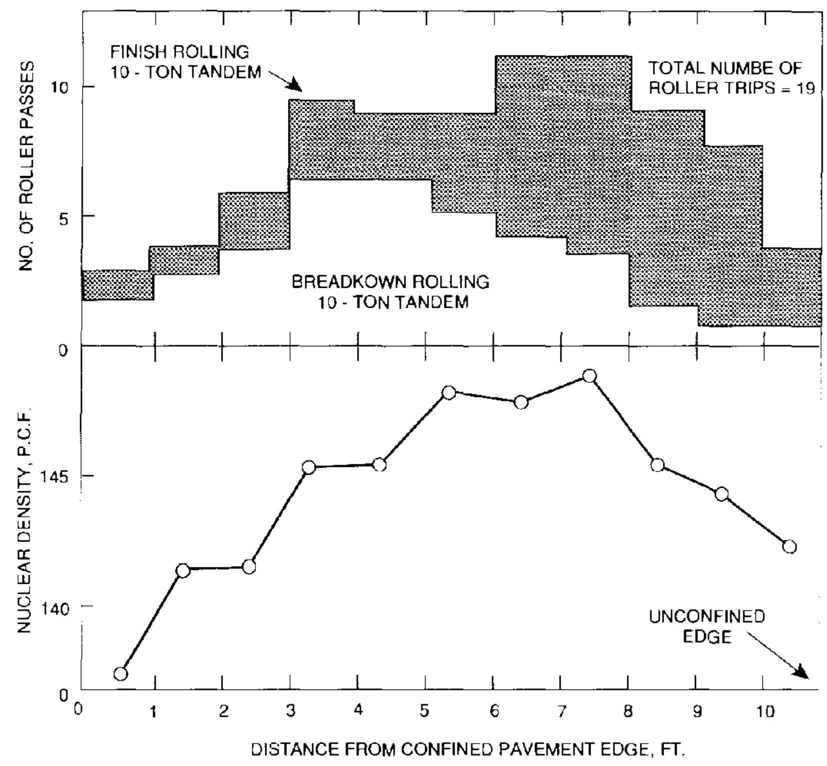


Figure 10.—Lateral variation in compactive effort and density.

Marshall Test Results

A number of State highway departments use the Marshall test and equipment to design the job-mix formula and control the ideal blend of aggregate, aggregate sizes, and bitumen, so that the mixture will be stable and durable when it is incorporated into the pavement. Marshall test data variations for stability, flow, and air voids from several State projects are shown in table 9. Testing and sampling variances for *stability* and *flow* values total 58 and 76 percent respectively. Variability of Marshall stability is shown in figure 11 for 3 groups of jobs: the third with the least variable standard deviation, the middle third, and the third with the most variable standard deviation. The computed σ̄'s from these groups were used to plot the normal curves.

Temperature

Another physical characteristic of the mix that may effect final density is mix temperature during breakdown rolling. Kilpatrick and McQuate (*8*) concluded that: "Breakdown rolling, both steel and pneumatic, should be completed before the pavement temperature drops below 220° F. to achieve maximum density." It is probable that final rolling, when accomplished above this critical temperature, will also produce the best results. The average standard deviation of temperature at the paver for 10 research jobs was 15° F. the range was from 6° to 22° F. Consequently, a plant producing batches with an average temperature of 275° F. will have a number of batches in the 230°-250° F. range. With temperatures in this range, it is difficult to achieve proper breakdown before the pavement cools below the reported critical 220° F.

Pavement Thickness

Thickness is another attribute needed to achieve economy of construction. A pavement that is thicker than required for adequate performance needlessly increases cost. A pavement that is too thin reduces service life and increases maintenance cost. Ideally, design thickness for pavements can be used to provide the most economical construction. However, research indicates that the variations in thickness of presently constructed pavements may significantly influence such performance. Data from 12 jobs in four States show that accepted surface courses have a σ̄ of 0.26 inch. In other words, about 5 percent of the pavement will have a thickness over ½ inch less than desired if the average corresponds to the specification. If concepts developed by Rex (9) are utilized, the computed expected service life for 5 percent of the area of a 3-inch pavement will be only ⅔ of the design life.

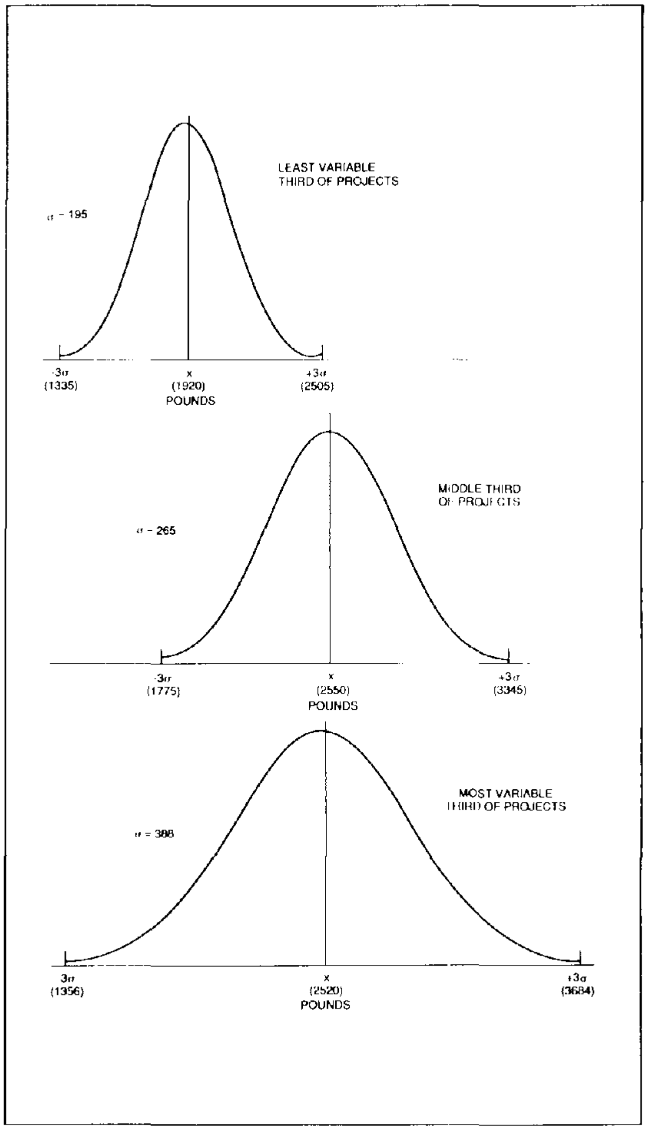


Figure 11.—Computed 3σ̄ limits of Marshall stability for 18 projects grouped according to size of average standard deviation, σ̄

Table 9.—Hot mix Marshall test data variations for stability, flow, air voids

	Projects	States	Average standard deviation (σ̄)	Average ($\bar{X}$)	Average variance components as a percent of total variance (σ_o^2)		
	Number	*Number*					
Marshal stability . pounds .	18	4	283	2,305	38	20	42
Marshall flow . . . 100/in . . .	15	2	1.29	8.62	62	14	24
Marshall air voids . . . pct . .	18	4	1.00	4.33	21	24	55

Conclusions

The production of high quality bituminous pavements requires the diligence of all concerned—the producer, the contractor and the contracting agency. The statistically measured variations (parameters) of accepted construction presented in this article indicate that much more variability exists than is revealed by the usual acceptance tests. Variations in excess of those normally expected for *good practice* were prevalent on almost every job studied. At present, the full significance of such variations cannot be assessed. Large sampling and testing errors virtually prevent a true evaluation of the material variation on a specific job. Also, it is difficult to assess the degree to which the variations affect actual pavement performance.

Because performance has not always been satisfactory, the need for improvement is obvious. Research results indicate that much improvement could be obtained and testing load reduced by the following changes:

- Adjust tolerance limits on gradation to conform to the principle of most tolerance on largest fraction retained on a sieve.
- Control the uniformity of gradation of the mixture by hot bin sieve tests, when a printed record of batch weights is available.
- Reduce to a minimum the number of sieves used for control testing.
- Exercise more diligence in the training and surveillance of operators performing control and acceptance tests.
- Require installation of automatic features on asphalt plants and finishers to reduce human error.
- Use random sampling to obtain all test portions.

REFERENCES

(*1*) *The AASHO Road Test Report No. 2, Material and Construction,* National Academy of Sciences, National Research Council, Publication 951, Highway Research Council, Publication 951, Highway Research Board Special Report 61–B, 1962, table 46, p. 82.

(*2*) *Guide Specifications for Highway Construction,* American Association of State Highway Officials, 1968.

(*3*) *The WASHO Road Test, Part 1: Design, Construction, and Procedures,* Highway Research Board Special Report 18, 1954.

(*4*) *A Study in the General Field of Quality Control Engineering,* State Road Department of Florida in cooperation with the Bureau of Public Roads, by Warden Engineers, Inc., Raleigh, N.C., September 1965.

(*5*) *A Statistical Study of Asphaltic-Concrete,* by Kenneth C. Afferton, New Jersey State Department of Transportation, Highway Research Record 184, 1967, pp. 13–24.

(*6*) *Determination of Statistical Parameters for Highway Construction,* Research Project No. 18, State Road Commission of West Virginia, by Materials Research and Development Inc., Miller Warden Associates Div., Raleigh, N.C., July 1966.

(*7*) *Realistic Job-Mix Formula Tolerances for Asphalt Concrete,* Research Report 66–3, by Malcolm D. Graham, William C. Burnett, Jerome J. Thomas, and Charles B. Fletcher, Bureau of Physical Research, New York Department of Public Works, January 1967, p. 21.

(*8*) *Bituminous Pavement Construction,* by M. J. Kilpatrick and R. G. McQuate, Bureau of Public Roads, June 1967, pp. 19–22, and figure 16, p. 23.

(*9*) *Effect on the Road of Variations in Thickness,* M-W Technical Paper No. 87, by Harry M. Rex, Miller Warden Associates, Raleigh, N.C., February 1964.

(*10*) *Procedures for Using Statistical Methods for Process Control and Acceptance of Bituminous Mixtures,* South Carolina State Highway Department, by Paguette-Mills, Consulting Engineers, Atlanta, Ga., February 1966.

(*11*) *A Statistical Analysis of Type 3 Plant Mix Bituminous Surfacing for Quality Control,* by Thomas P. Kostohris, State of Montana Highway Commission, 1966.

(*12*) *Bituminous Concrete Shoulder Study, Part 2—Analysis of Depth Variation Related to Two Base Course Materials,* by Fred E. Sternberg, Connecticut State Highway Department, November 1966, pp. 13a–22a.

(*13*) *Statistical Approach to the Quality Control of Plant Mix Pavement,* Wyoming State Highway Department, June 1968.

(*14*) *Statistical Parameters Research Quality Control Study on Asphalt Pavement,* Colorado Department of Highways, April 1967.

(*15*) *A Study of Variability in Production, Sampling and Testing Bituminous Concrete Base, Part II—the 1966–67 Study,* by Richard French and Bryce Edwards, Maine State Highway Commission, September 1967.

(*16*) *Tolerances for Asphaltic Concrete and Base Courses,* by Verdi Adam and S. C. Shah, Louisiana Department of Highways, November 1966.

(*17*) *Variations in Bituminous Construction in Nebraska,* Nebraska Department of Roads in cooperation with the Bureau of Public Roads, September 1968.

(*18*) *Bituminous Concrete Surface Course Analysis,* State Highway Commission of Wisconsin in cooperation with the Bureau of Public Roads, 1965.

(*19*) *Quality Control of Construction by Statistical Tolerances,* by J. H. David, Alabama Highway Department, May 1967.

(*20*) *Variability in an Asphalt Concrete Mix,* by Edward R. Oglio and Joseph A. Zenewitz, PUBLIC ROADS, A JOURNAL OF HIGHWAY RESEARCH, VOL. 34, No. 1, April 1966, pp. 5–15.

(*21*) *Compaction of Bituminous Pavement and Base Courses,* Production Study Report No. 401, by Edwin C. Granley and Richard G. McQuate, Bureau of Public Roads, U.S. Department of Commerce, July 1965.

Part 5—

Summary of Research for Quality Assurance of Aggregate

Reported by JAMES A. KELLEY,
Highway Research Engineer,
Materials Division

Introduction

A review of the evaluation by statistical techniques of highway aggregate characteristics is presented here as a condensed compilation of both historical data and data from designed quality-measurement projects in which the degree of conformance to specifications was statistically estimated. The historical data are not sufficient to determine the reason for any nonconformance to the specifications. However, the designed quality-measurement projects do provide data to determine quality at any point in a process, to disclose operations needing corrective action, and to give a valid estimate of specification conformance.

Reports from nine States on projects in which research data have been obtained are abstracted and summarized in this compilation to illustrate trends in gradation analysis. sampling and testing procedures, sand equivalent analysis as an alternate to gradation analysis, and soundness tests for aggregate quality.

Aggregate Base Course Characteristics

Specifications for base course aggregate usually contain limits for gradation, plasticity, soundness, and amount of deleterious material. Variations in gradation have been studied rather extensively to ascertain the degree of conformance obtained in construction. The data have been analyzed statistically to determine the variation in the material itself and that arising from sampling and testing. Most of the studies have been projects sponsored cooperatively by Public Roads and State Highway Departments, although some bave been entirely State funded.

Nonuniformity of the final product has been disclosed by results of studies of gradation of different aggregate types including gravel, sand-gravel, and crushed stone. Differences in gradation were found between samples taken from the borrow pit or quarry plants, from the material after stockpiling, and again, from the material after it had been processed and compacted in place on the roadway. Differences in test results on the aggregate often resulted from the sampling method—sampling from a moving or stopped belt compared with sampling from a loaded truck. Representative sampling from an operation or placement also gave results that differed from those obtained by random sampling.

Combined variations frequently add up to a total variance of such magnitude that assurance of compliance with specifications is doubtful. However, with the knowledge provided by statistical analysis, it has been possible not only to pinpoint areas or operations requiring improvements, but also to determine when to take immediate corrective measures to assure better compliance.

Variance in historical data

Early statistical studies were made on data in office files of completed projects. Although this type of data was not randomly selected, statistical analysis usually disclosed that measurements of base course characteristics followed a normal distribution.

In table 1, which was extracted from a study of historical data for 257 observations of type A base in Louisiana, it is shown that for projects considered acceptable, the mean of the distribution for all sieve sizes was well within design limits. However, the statistically computed percentage of material within the design limits varied for each sieve size. The lowest value was 82 percent for material passing the No. 40 sieve. The highest value was 99 percent for material passing the ¾-inch sieve.

Variance of controlled research data

In the State of West Virginia, new construction was evaluated statistically to determine variations from design gradations. Analysis of variance disclosed that the material variance tended to be large and the sampling and testing variances small. According to the data, the magnitude of variance seemed to be directly related to the amount on each sieve.

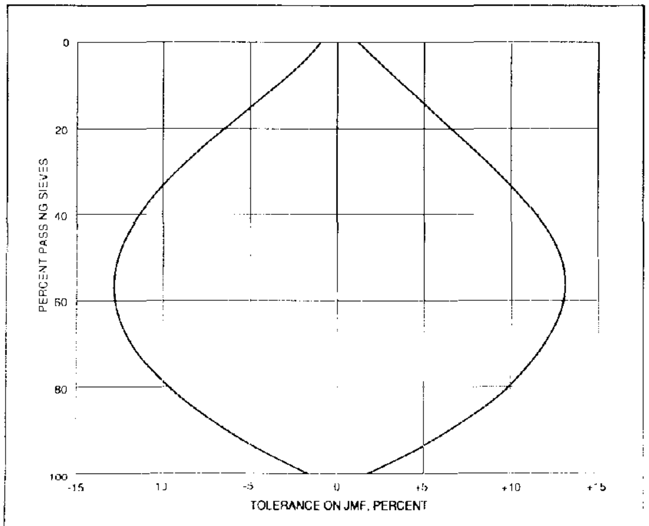


Figure 1.—Aggregate base course gradation characteristics, 95-percent probability tolerances on job-mix formula, West Virginia, 1966.

The data in table 2 are an example of many studies in West Virginia and other States in which the components of variance are isolated by statistical analysis of field data on aggregate gradation characteristics. In figure 1, which is a diagram from the West Virginia report, proposed 95 percent tolerance limits are shown. The tolerances are ±13 percent on the sieve having approximately 50 percent of the material passing, and taper in both directions toward 0 percent and 100 percent passing where the tolerances are ±2 percent.

Variance caused by operators, sampling methods and equipment

Variance in the gradation of aggregate mixtures often is the result of sampling and testing procedures, as well as of the material itself. Several States have made quantitative measurements of these parameters. In Michigan (*1*)[1] a field experiment was carried out to determine what part aggregate inspectors, screening sieves, and sampling methods play in the uniformity of gradation results. A mathematical model was prepared to analyze the variations and ascertain whether (1) inspectors require further training to sample and test aggregates, (2) testing equipment requires periodic calibration or

---

[1] Italic numbers in parentheses identify the references listed on p. 54.

Table 1.—Base course analysis, gradation type A—historical data, Louisiana

Sieve size	Design limits	Mean distribution (x̄)	Standard deviation (σ)	Compliance with design
	Pct.	*Pct.*	*Pct.*	*Pct.*
¾ in.	75–95	90	2.5	99
No. 4	40–60	55	4.9	91
No. 40	20–45	37	6.3	82
No. 200	10–20	16	2.9	91

Table 2.—Base course gradation analysis—research data (n = 136), West Virginia

Sieve sieve	Design limits	Mean distribution (X̄)	Standard deviation (σ)	Variance		
				Material (σ_a^2)	Sampling (σ_s^2)	Testing (σ_t^2)
	Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
1½ in	100	100	0.0	0.0	0.0	0.0
¾ in	40–85	80	3.9	9.6	5.8	0.2
⅜ in	....	50	5.2	18.3	6.4	2.4
No. 4	20–60	34	4.3	12.7	4.9	1.1
No. 16	....	20	3.6	9.1	3.1	0.6
No. 40	5–25	11	2.8	5.7	2.6	0.0
No. 100	....	6	2.7	4.8	2.9	0.0

maintenance, (3) improved precision is feasible in gradation analysis, and (4) significant interactions occur in the experimental work. The results of this study were as follows:

- Individual inspectors and methods of sampling had a relatively small effect on gradation results on the ⅜-inch sieve. According to an analysis of components of variance, an estimated 4 percent of the total variance was attributable to inspectors, 6 percent to sampling methods, and the remaining 90 percent to inherent material and experimental deviations.
- For material passing the No. 10 sieve, significant interaction effects among the main factors of the experiment were shown to exist. Variance of 0–8 percent was due to methods of selecting samples, variance of 7–18 percent was due to testing and the remaining variance was attributable to inherent material and experimental deviations.
- The results of the analysis of variance (see table 3) indicated that interaction effect was significant enough to reduce the accuracy of major comparisons. According to the data in table 3, the combined influence (interaction) of inspectors and screening kits affected the gradation results. Also, the State found that the difference between the two sampling methods was large enough to be of practical importance. The relative performance of aggregate inspectors was not consistent for all screening kits. These variances were significant, although not as large as the material variance, and it was presumed that, with training and corrective maintenance, the amount of testing and sampling variance could be reduced.

Methods of automatic aggregate sampling from a belt delivery system, and the variance resulting from the method used to prepare the test sample were studied in Idaho. Samples obtained with an automatic sampling device produced lower variance than those obtained manually, and the variance was more uniform. A direct relation was found between the splitting method and the testing variance of samples. Cross-split samples had a lower variance than those split only once. Cross-splitting is similar to quartering on a mat and combining the opposite quarters to form a single sample. Researchers tested 34 samples from Pit Le–111, collected by the manual method, and 25 samples from Pit Jr–2, obtained with an automatic sampling device. The variances for Jr 2 are relatively small and much more uniform than those for Pit Le–111. Part of the difference was attributed to the difference in the splitting techniques. The Idaho report was prepared to permit several cross comparisons of testing and sampling work. On the basis of these tests, 17 percent of the overall variance was due to testing variance whereas 30 and 53 percent, respectively, were due to sampling and material variances.

In Idaho, extensive research (*2*) was also conducted to ascertain whether the sand-equivalent test procedure was sufficiently reproducible to determine aggregate acceptability. The tests performed on cross-split samples at the Moscow laboratory resulted in a testing variance of 0.96, whereas the single-split samples at the Boise laboratory resulted in a testing variance of 1.85. For sand-equivalent determinations, considerable discrepancy existed between the results of the two laboratories; however, the test was considered satisfactory if the cross-split technique of the Moscow laboratory was used. As a result of the statistical analysis, improvements in both sampling and testing methods were initiated.

A study in California (*3*), was undertaken to evaluate the effectiveness and reliability of the sand-equivalent tests used for procedure control and for measuring the variation of the aggregate investigated. Tests were performed on 200 random samples from each of six projects. Gradation was determined for each sample, and the analysis of variance was reported for the results on several sieves. It was concluded that the sand-equivalent and sieve analyses, supplemented by the R-value results in borderline situations, can provide satisfactory control of base and subbase material. The variances for the test results on the base material were generally smaller than the

Table 3.—Analysis of variance for passing No. 10 sieve, Michigan [1]

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]	97.98	1	97.98	[3]10.67	3.90	6.81
	I [4]	39.10	2	19.55	2.12	3.06	4.75
	S [5]	14.06	2	7.03	0.77	3.06	4.75
Interactions among factors	MI [2 4]	25.29	2	12.64	1.38	3.06	4.75
	MS [2 5]	3.61	2	1.80	0.20	3.06	4.75
	IS [4 5]	280.20	4	70.05	[3]7.63	2.43	3.45
	MIS [2 4 5]	98.81	4	24.70	[6]2.69	2.43	3.45
Replication	Residual	1,487.32	162	9.18	....	....	....
	Total	2,046.36	179	11.43	....	....	....

[1] Michigan Report No. R–571.

[2] M Sampling methods.

[3] Significant at the 1 and 5-percent levels (highly significant).

[4] I Aggregate inspectors.

[5] S Screening kits.

[6] Significant at the 5-percent level.

variance for the subbase material, perhaps because of the greater selectivity used for base material. Although the sampling and testing variances were relatively small for both materials, the testing variance was significantly larger than the sampling variance.

The results of this research were used to propose revision of California aggregate specifications. The proposed revisions, shown in table 4, were designed so that present specification limits could be retained by basing acceptance on a moving average of the five most recent test results. Broader limits for individual test results were established. Based on information available to him, the resident engineer is now authorized to accept the material, provided that the average indicates that the process is in control, even though a single test result may deviate from the broader limits.

According to the California report, class 2 aggregate base had an average sand-equivalent value of 44 with a pooled standard deviation of 4.8, and class 2 aggregate subbase had an average sand-equivalent value of 32 with a pooled standard deviation of 5.0. The proposed specification requirements for the sand-equivalent test and gradation are shown in table 4. It was stated in the report that:

"... the proposed specifications are to be used as guidelines only and are not intended to interfere with the present practice of designing specifications to meet local conditions for economic reasons. Once the gradation limits are established for a particular job, statistical specifications can be designed using the standard deviation as reported in this study, if no more accurate measurements are available."

After publication of the report, the State Division of Highways used similar specifications in its construction of projects.

Salt soundness test of aggregate

In certain uses, the quality of individual aggregate particles is an important characteristic, and owing to the composition of gravel or stone, the soundness of the aggregate pieces must be determined by certain standardized tests. In a study of *Salt Soundness Tests for Fine Aggregate* (*4*), the New York Department of Transportation used statistical concepts to investigate the procedures for determining both the sodium and magnesium salt soundness of fine aggregate and the methods used to judge the acceptability of a source. Data were presented on (1) the effect of drying time on the magnitude and reproducibility of test results, (2) overall reproducibility of the test with sodium and magnesium sulphates, and (3) the combined effect of testing and production variations on the scatter of test results from single sources. The summary statistics for the soundness tests, with various drying periods, is shown in table 5.

The conclusions extracted from the New York study were "(1) that an increase in drying time in the test from 6 to 30 hours will result in no change in the magnitude or reproducibility of the test results, (2) the reproducibility of the test with sodium sulphate and the test with magnesium sulfate are not significantly different, and (3) that it is possible to place the acceptance of sources of fine aggregate on a sound statistical foundation." They recommended that "the magnesium sulphate soundness test continue to be performed at the rate of one cycle per day and that the test with sodium sulphate be discontinued."

Even though the results of the New York study of fine aggregates were generally acceptable as reproducible results, many States have not obtained satisfactory correlation between salt-soundness test results and performance. This is particularly true for coarse aggregates.

Table 4.—Digest of proposed specifications for class 2 base and subbase aggregates, California

Material	Sand-equivalent values (Test Method, California 217)			Gradation values		
					Percent passing	
	Minimum average [1]	Not to be lower than [2]	Overall average [3]	Sieve size	Moving average	Individual test result
Base	30	25	36	1 inch	....	100
				¾ inch	95 ± 5	95 (+5)(−7)
				No. 4	45 ± 10	45 ± 15
				No. 30	20 ± 10	20 ± 13
				No. 200	5.5 ± 3.5	5.5 ± 4.5
Subbase	23	18	30	3 inch	....	100
				2½ inch	95 + 5	95 (+5)(−10)
				No. 4	65 ± 25	65 (± 35)
				No. 200	12.5 ± 12.5	12.5 (+17.5)(−12.5)

[1] Five consecutive tests, each performed on independent sample.

[2] No single sand equivalent result to be lower.

[3] Overall average should be maintained for 99.9 percent probability of acceptance of suitable material.

Table 5.—Summary statistics for magnesium sulfate soundness tests, New York

Sand No.	Arithmetic means [1]				Variance estimates [1]			
	Drying period		Difference in variation [2]	Higher	Drying period		Difference in variation [2]	Higher
	6-hour	30-hour			6-hour	30-hour		
	Pct.	*Pct.*			*Pct.*	*Pct.*		
1	5.64	5.90	Insignificant	. . . .	0.23	0.24	Insignificant	
2	17.09	15.74	. . . . . do	. . . .	0.05	1.83	Significant	30-hour.
3	23.21	23.83	Significant	30-hour . . .	0.06	0.24	Insignificant	
4	47.01	41.98	. . . . . do	6-hour	1.37	2.88	. . . . . . do	
5	47.65	50.36	. . . . . do	30-hour . . .	2.13	0.40	Significant	6-hour.

[1] Calculated from results of tests on two groups of three samples each.
[2] Statistical significance at 0.05 confidence level.

Table 6.—Results for surface mixture samples, South Carolina

Sieve size	Sample location	Specification limits percent passing	Control chart values		Standard deviation		Analysis of variance			
			Average $(\bar{x})$	Total (σ)	Average $(\bar{x})$	Total (σ)	Total (σ_{to}^2)	Material (σ_m^2)	Sampling (σ_s^2)	Testing (σ_t^2)
		Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
½ in	Plant [1]	87–97	92.0	2.88	91.8	3.10	9.62	2.45	0.0	7.60
	Spreader [1]		90.0	4.00	90.7	3.52	12.41	8.74	0.0	3.50
	Compacted [1] . . .		92.2	3.30	92.1	3.16	10.02	0.0	2.18	6.19
No. 4	Plant	58–72	66.1	3.90	66.8	3.92	15.37	9.70	0.0	5.15
	Spreader		65.2	5.71	65.7	5.84	34.06	30.9	0.0	4.75
	Compacted		65.0	4.32	65.2	4.28	18.34	0.0	10.18	8.11
No. 10	Plant	42–58	52.0	3.45	52.6	3.69	13.59	8.33	1.76	3.69
	Spreader		53.0	4.98	52.7	5.60	31.35	28.15	0.0	5.04
	Compacted		54.3	3.92	54.5	4.01	16.11	0.0	11.36	7.39
No. 40	Plant	21–35	28.1	1.70	28.3	1.91	3.64	2.28	0.0	1.40
	Spreader		28.0	1.41	28.5	2.34	5.48	4.92	0.0	1.00
	Compacted		28.7	1.85	28.8	2.08	4.31	0.0	2.51	2.09
No. 200	Plant	4–10	5.7	1.06	5.8	1.15	1.32	1.15	0.0	0.18
	Spreader		5.7	1.20	6.34	1.38	1.91	0.0	0.0	1.44
	Compacted		6.5	1.04	6.4	1.03	1.05	0.30	0.48	0.20

[1] Number of tests performed: Plant = 40, Spreader = 24, Compacted = 128.

Bituminous Concrete Aggregate Characteristics

Aggregate used in bituminous concrete mixtures is subjected to several manipulations and treatments that are not applied to base course aggregate. The aggregate is heated for drying and mixing with asphalt. Often, it is stockpiled or placed in storage bins before the mixing operations. The final mixture is spread by a mechanical spreader and then a high force is applied for final compaction and rolling. Thus, the finished layer has experienced many abrasive forces that could cause not only changes in gradation of the aggregate component, but also changes in density and stability. A more detailed analysis of variations in aggregates used in bituminous construction is contained in Part 4. However, the more important findings of individual projects are reported here.

In a study performed in South Carolina (5), random samples of asphalt mixtures were selected from trucks at the batch plant, from the roadway just behind the spreader, and from the roadway after compaction, to determine whether any progressive change occurred in the characteristics of the aggregate. A summary of this work is given in tables 6 and 7 in which the specification limits and analysis of variance for both surface and binder courses are also shown. The aggregate passing the No. 4 sieve in the surfacing mixture was within the job-mix formula only 50 percent of the time by routine control sampling and 66 percent of the time by random sampling. The material passing the No. 40 sieve was within the job-mix formula 76 percent of the time by control sampling and 88 percent of the time by random sampling. The test results shown in table 6 indicate that the average for the No. 4 material varied from 66.8 to 65.2 percent whereas the No. 40 material varied only from 28.3 to 28.8 percent, conforming more closely to the job-mix formula. The greatest standard deviation occurred on the samples from the spreader box.

The characteristics of aggregate used in bituminous mixtures were also explored in West Virginia (6). An analysis of the aggregate passing the No. 4 sieve is shown in table 8 for 10 bituminous projects. For the percent passing the No. 4 sieve, nine of the 10 projects had an average value that was within the specifications. However, the overall standard deviation for the individual projects was so large that many of the projects had a considerable amount of nonconforming material. Because of the large overall standard deviation, 4.4 percent, a change in the specificied job-mix tolerances was recommended. The following excerpts were taken from the West Virginia report:

"Tolerances for percentages passing other sieves may also require adjustment. Inspection of the data shows that the major component of the overall standard deviation, σ_o, is the material variance, σ_a, and sampling and testing can be reduced to a negligible amount."

"The size of the standard deviation of the percent passing any sieve, neglecting sampling and testing error, depends

Table 7.—Results for binder mixture samples, South Carolina

Sieve size	Sample location	Specification limits percent passing	Control chart values		Standard deviation		Analysis of variance			
			Average ($\bar{x}$)	Total (σ)	Average ($\bar{x}$)	Total (σ)	Total (σ_{to}^2)	Material (σ_m^2)	Sampling (σ_s^2)	Testing (σ_t^2)
		Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*	*Pct.*
1 in	Plant [1]	80–97	93.4	3.95	93.9	4.13	17.09	0.0	2.60	13.04
	Spreader [1]		93.8	5.19	93.3	4.07	16.56	3.35	0.0	12.25
	Compacted [1] ...		93.4	4.70	93.7	4.48	20.07	4.59	0.0	15.43
No. 4	Plant	35–50	40.6	4.76	40.8	4.61	21.25	8.47	7.36	7.52
	Spreader		40.1	6.79	41.3	6.01	36.14	21.20	5.37	12.87
	Compacted		43.1	4.42	42.9	4.54	20.60	4.23	0.0	8.67
No. 10	Plant	25–35	32.2	4.06	32.2	3.90	15.24	7.22	3.68	4.43
	Spreader		32.2	5.46	32.9	4.83	23.3	14.03	0.0	8.38
	Compacted		34.8	3.51	34.6	3.74	13.96	0.0	5.86	5.91
No. 40	Plant	None	18.0	2.09	18.1	2.18	4.74	2.17	1.14	1.46
	Spreader		18.3	2.66	18.5	2.34	5.48	3.18	0.0	2.37
	Compacted		19.7	2.00	19.9	3.62	13.15	2.24	0.0	10.87
No. 200	Plant	None	4.1	0.53	4.1	0.54	0.30	0.17	0.05	0.09
	Spreader		4.0	0.49	4.1	0.58	0.33	0.18	0.0	0.18
	Compacted		4.3	0.67	4.3	0.67	0.44	0.11	0.18	0.15

[1] Number of test performed: Plant = 284, Spreader = 68, Compacted = 380.

Table 8.—Analysis of variance of bituminous concrete aggregate for 10 projects in West Virginia

Project No.	Sample location	Number of samples (n)	Percent passing No. 4 sieve Specification 60–70				
			Average ($\bar{X}$)	Overall standard deviation (σ_o)	Standard deviation		
					Material (σ_a)	Testing (σ_t)	Sampling (σ_s)
			Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*
38 A1	Even	96	66.7	2.5	2.5	0.0	0.0
38 A1	Odd	96	67.5	4.8	2.5	1.6	3.8
3235	Even	100	65.6	3.6	3.6	0.0	0.0
3235	Odd	100	64.8	3.3	3.3	0.0	0.0
3462	Truck	120	69.3	4.5	3.4	0.0	2.9
3462	Pavement	120	65.7	4.7	4.1	0.8	2.1
173 H(1) & (2)	....	200	70.0	4.8	3.6	1.5	2.2
204A & 204A (3)	....	180	72.1	5.1	4.4	1.8	1.7
284(C) & (4) SRC	....	120	61.7	4.0	3.9	0.8	0.3
284(C) & (4) AASHO	....	120	61.5	4.2	3.9	1.6	0.5
Average, all projects	....	....	66.5	4.4	3.7	1.2	1.9

to a large extent upon the value of the percentage passing that sieve" (fig. 2).

In the West Virginia report it was proposed that tolerances for gradation specifications be varied according to the percentage passing any sieve. The magnitude of variation to provide 95 percent probability tolerances on the job-mix formula is shown in figure 3.

Other States engaged in statistical studies of aggregate-gradation characteristics in bituminous mixtures have indicated that, for best uniformity and smallest standard deviations, control of gradation should be at the mixing plant. Job-mix tolerances for all gradations should be adjusted for the percentages expected to pass the specified sieves.

Portland Cement Concrete Aggregate Characteristics

Because structural concrete in highway construction is critical, the specified aggregate gradation should be assured. Several research projects were conducted to determine the best place to sample aggregate for control, permissible tolerances on various sieve sizes, and alternate methods or tests to establish gradation uniformity.

In California, a study was performed to determine the precision of current test methods and the feasibility of using statistical quality control procedures for portland cement concrete aggregate. Several conclusions were drawn from this study: Present controls and specifications for aggregate gradation need to be modified because of the high material variance and large percentages of out-of-specification gradation; sand-equivalent and cleanliness test methods were satisfactory; more efficient field control would be possible if control charts were used; better control of gradation could be obtained by using a *moving average* based on the results of the five most recent individual tests; material and testing

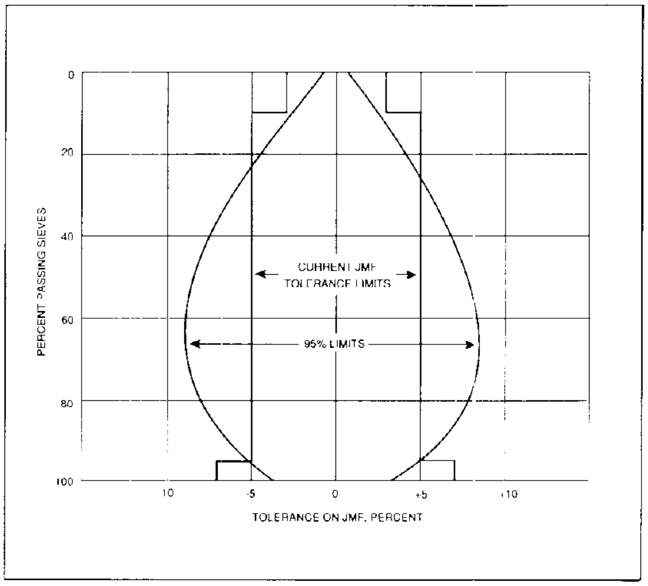


Figure 3.—Relation of 95-percent probability tolerances on job-mix formula to percent passing sieves, asphaltic wearing course, West Virginia, 1966.

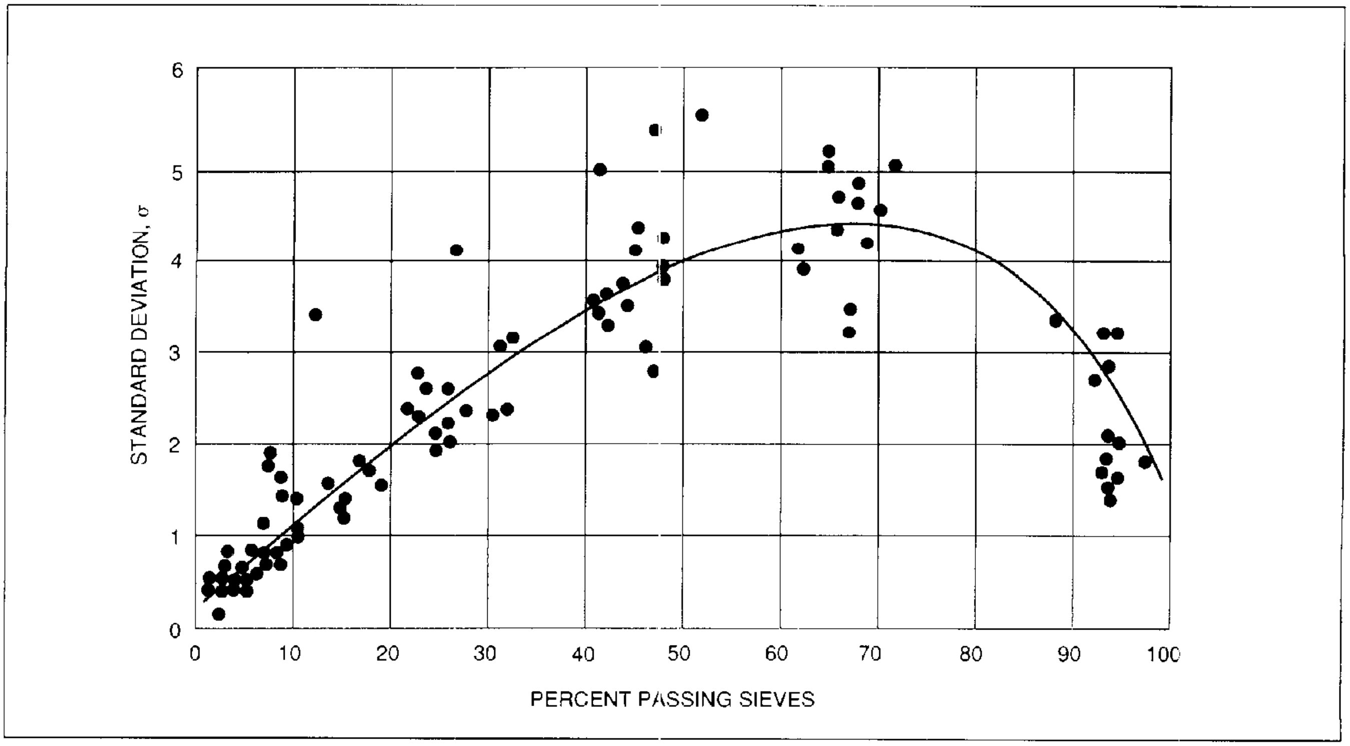


Figure 2.—Relation of standard deviation to percentages passing sieves, asphaltic concrete wearing course, West Virginia, 1966.

variances were considerably larger than were anticipated (see fig. 4); and a relatively high percentage of the aggregate failed to meet the specification, which is shown by the diagram in figure 5.

A statistical analysis of variance in aggregate for portland cement concrete was made by Louisiana (*7*). The variations in gradation of fine and coarse aggregate sampled from different stockpiles as well as the differences between samples with stockpiles, were determined. According to the Louisiana report, "The largest component of variance is between stockpiles, which is reflective of material variance. The variation between samples within stockpiles can be attributed to either the stockpiling technique or sampling procedure." The actual results for the fine aggregate passing the No. 4 sieve are shown in table 9. The analysis of the coarse aggregate was similar to that of the fine aggregate. As shown in table 10, heavily loaded sieves had the greatest deviations and the largest amounts of material outside the specification limits. As a result of this study, the researchers prepared suggested acceptance limits and frequencies of measurement for aggregate used in portland cement. (See table 11.)

Quality control of aggregate used in portland cement concrete by sampling from the stockpiles and bins at the central plant was studied in Oklahoma (*8*). The dry aggregate

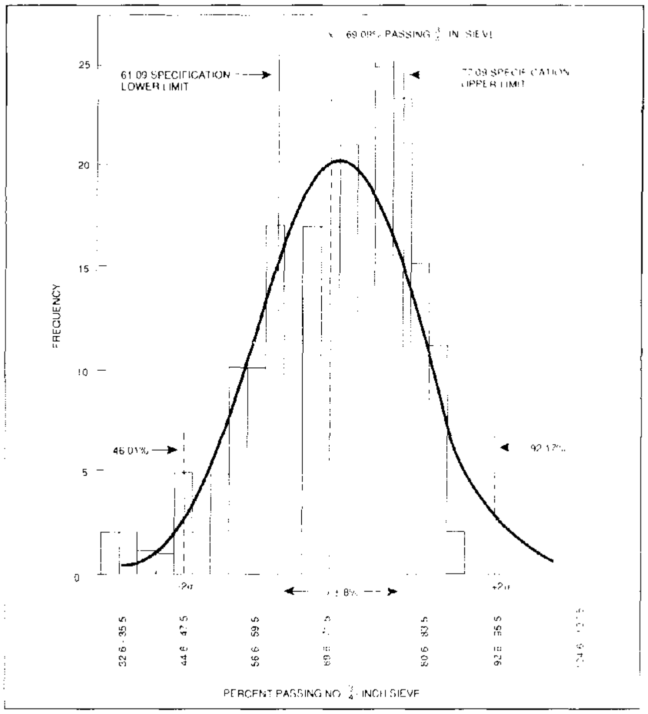


Figure 5.—Percent of material outside of specifications, portland-cement-concrete aggregate passing No. ¾-in. sieve, project No. 2, California.

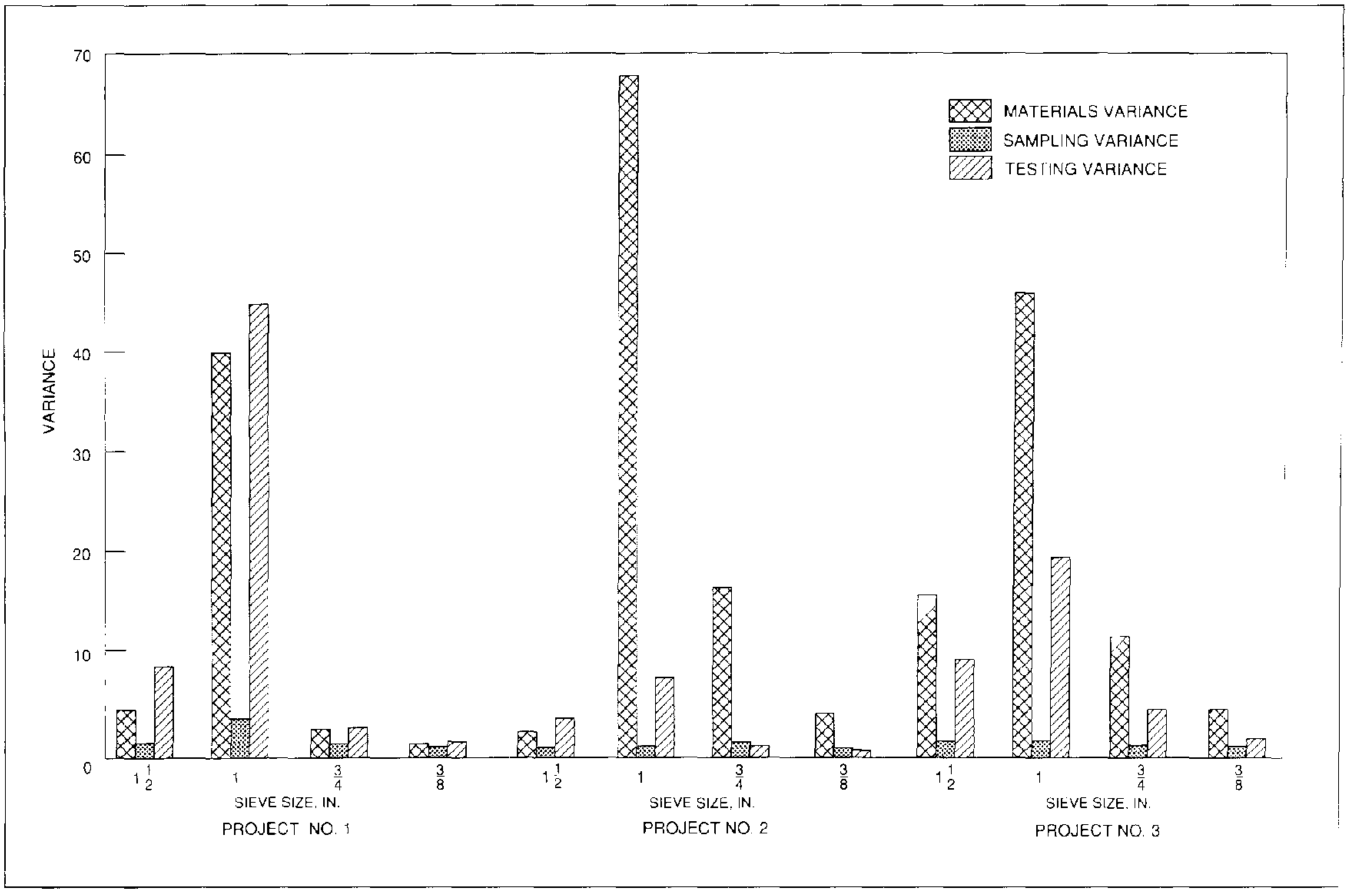


Figure 4.—Analysis of variance of portland-cement-concrete coarse aggregate for material passing different sieves, California.

Table 9.—Analysis of variance on gradation of fine aggregate for portland cement, Louisiana, Percent passing No. 4 sieve

Source of variance	Sum of squares (SS)	Degrees of freedom (DF)	Mean squares (MS)	Estimate of mean squares [1] (EMS)	F.05
Between stockpiles	249.43	8	31.18	$\sigma_0^2+\sigma_s^2+16\sigma_{st}^2$	[2] 8, 63
Between samples within stockpiles	52.44	63	0.83	$\sigma_0^2+2\sigma_s^2$	[2] 63, 72
Between subsamples within samples	14.35	72	0.20	σ_a^2	....
Total	316.22	143	....	....	....

[1] σ_e^2 = .20 (Testing) σ^2 = .32 (Sampling) σ^2 stockpile = 1.90 (Material).

[2] Significant

Table 10.—Summary of statistical results on portland cement concrete aggregate gradation, Louisiana[1]

Sieve size	Average ($\bar{x}$)	Standard deviation (σ)	Minimum	Maximum	Outside specifications	Specification limits	Variance (σ^2)		
							Test	Sample	Stockpile
Grade A course aggregate									
	Pct.	*Pct.*	*Pct.*	*Pct.*	*Pct.*		*Pct.*	*Pct.*	*Pct.*
1 in	95.6	3.8	82.7	99.9	7.3	90–100	0.84	5.65	9.97
¾ in	75.4	10.8	46.1	88.8	2.1	40–88	6.41	77.62	40.72
½ in	35.5	12.7	4.6	60.2	13.5	15–55	9.18	132.04	26.72
No. 4	1.3	1.2	0.2	5.5	0.0	0–6	0.13	0.76	0.72
Fine aggregate									
No. 4	97.8	1.5	92.1	99.9	3.8	95–100	0.20	0.32	1.90
No. 16	79.2	7.9	56.6	91.6	9.7	45–90	0.70	10.72	57.69
No. 50	15.9	6.5	7.2	31.6	1.4	7–30	5.36	39.88	0.0
No. 100	2.1	1.3	0.3	5.7	0.0	0–7	0.04	0.40	1.34

[1] Louisiana Department of Highways Report, 1966.

Table 11.—Suggested acceptance limits for portland cement concrete aggregate [1]

Sieve size	Acceptance probability (Pa)	Rejection probability (Pr)	*n*	Acceptance limits				Measurement frequency
				Mean		Individual		
				LL	UL	$\bar{x}s-$	$\bar{x}s+$	
Gradation of fine aggregate, percent passing								
	Pct.	*Pct.*						
No. 4	99	90	4	95.90	99.75	3.93	3.93	One every 200 cu. yd.
No. 16	99	90	4	68.97	89.45	20.80	20.80	
No. 50	99	90	4	7.51	24.35	17.11	17.11	
No. 100	99	90	4	0.41	3.71	3.35	3.35	
Gradation of grade A coarse aggregate, percent passing								
1 in	99	90	4	90.63	100.00	10.10	10.10	One every 500 cu. yd.
¾ in	99	90	4	61.43	89.39	28.40	28.40	
½ in	99	90	4	19.08	51.84	33.27	33.27	
No. 4	99	90	4	0.0	2.70	3.20	3.20	

[1] Louisiana Department of Highways Report, 1966.

was weighed at the bin site, the cement added, and the batch hauled by trucks to the road site where the concrete was mixed. Random samples were taken at a point in the stockpile nearest the bins. The analysis of the gradation indicated that the mean values for each sieve size were within the specification limits although many individual values were outside the upper and lower control limits. The Oklahoma Department of Highways recommended that the gradation determination be continued, but with certain modifications. Acknowledging that some plans provide for the sampling of aggregates at the batching bin, researchers pointed out in their report that sampling at the stockpiles permits early detection of undesirable or unacceptable aggregate, which is the purpose of quality control—to locate defective material as quickly as possible.

Summary

Some of the important findings from selected research on the characteristics of aggregate used in base courses and in bituminous and portland cement concrete mixtures have been presented here. More attention has been given to aggregate-gradation characteristics from source of supply to placement, than to other characteristics. Early studies concentrated on historical data; more recent studies were conducted during actual field construction. Comprehensive plans were devised to study historical data and to measure variability during construction. The degree of conformance to gradation specifications was found to vary from step to step in the processing. Analysis of variance usually was applied during construction to determine causes of the variation and to locate conditions needing corrective action.

Generally, the largest deviations from specifications were in the material in the middle of a stack of sieves—where a large amount of material is on individual sieves.

Knowledge of inherent material, sampling, and testing variations enables the engineer to design specifications with tolerances that are compatible with local conditions and, thereby, to avoid unenforceable requirement or unreasonable expense and still obtain a suitable aggregate.

For aggregate control, the sand-equivalent test rather than gradation is preferred in some States, and statistical research, conducted to ascertain wehther the sand-equivalent test is informative and reproducible, has confirmed its usefulness, although the amount of data at present is rather limited.

Statistical research on salt-soundness determination indicated that drying time did not need to be changed; that sodium sulfate testing could be discontinued, as magnesium sulfate testing is satisfactory; and that it is possible to place acceptance of fine aggregate sources on a sound stastistical foundation.

In statistically-oriented research on aggregate gradation for bituminous mixtures, information similar to that for base courses was developed, and indicated that variation from the specifications differs according to the point of sampling. The research showed that the variation on certain sieve sizes is considerable indicating either the need for improvement in sieving operations or the establishment of wide tolerances in the specifications to eliminate compliance disagreement. The research also indicated that sampling at the hot bins was preferred for process control, whereas, sampling at the compacted bituminous layer was best for establishing uniformity of the mixture.

Data in various studies indicated that the aggregate used in portland cement concrete had a smaller standard deviation than the aggregate used in bituminous mixtures or in base courses, but that statistical analysis provided information for early detection of undesirable gradation or undesirable quality.

Based on their studies, highway departments in some States are revising their specifications and outlining specific sampling procedures. The use of statistically designed control charts is highly recommended for control of the characteristics of aggregate by these States. The moving average of five most recent individual test results is reported to be the most practical for controlling the construction processes.

REFERENCES

(*1*) *Highway Quality Control Program, Statistical Parameters,* Research Report No. R–572, Michigan Department of State Highways, March 1966, p. 41

(*2*) *Quality Control,* Research Project No. 11, State of Idaho Department of Highways, May 1967, p. 16

(*3*) *A Statistical Analysis of Untreated Base and Subbase Materials,* State of California Transportation Agency, Department of Public Works, Division of Highways, March 1967, pp. 22 and 23.

(*4*) *Salt Soundness Tests for Fine Aggregates,* New York State Department of Public Works, September 1964, p. 3.

(*5*) *Procedures for Using Statistical Methods for Process Control and Acceptance of Bituminous Mixtures,* prepared by Paquette-Mills, Consulting Engineers, Atlanta, Ga., for South Carolina State Highway Department, February 1966, pp. 78 and 79.

(*6*) *Determination of Statistical Parameters for Highway Construction,* Research Project No. 18, prepared by Materials Research and Development, Inc., Miller-Warden Associates Div., for the State Road Commission of West Virginia, November 1966, p. III–49.

(*7*) *Quality Control Analysis, Part III, Concrete and Concrete Aggregates,* by S. C. Shah, Louisiana Department of Highways, November 1966, p. 11

(*8*) *Statistical Quality Control of Portland Cement Concrete Pavement,* prepared by Joakim G. Laguros, University of Oklahoma, for Oklahoma Department of Highways and U.S. Bureau of Public Roads, June 1968.

Part 6—
Control Charts

Reported by EDWIN C. GRANLEY,
Highway Research Engineer,
Materials Division

Introduction

Test measurements portrayed graphically demonstrate the familiar adage, *a picture is worth a thousand words*. One type of graphical portrayal, the control chart, is used extensively by industry in quality-assurance procedures. Many of the industrial control-chart principles are just as applicable to process control and acceptance inspection in highway construction as they are to industrial techniques.

Control charts show cumulative trends in dimensional or physical properties within maximum and minimum limits that denote acceptable production. They not only indicate when established limits have been exceeded, but also provide the means to anticipate and correct causes that tend to promote the production of defective products. Thus, the use of control charts is an application of the principle that an ounce of prevention is worth and a pound of cure.

In industrial statistical quality control methods, certain variations in product quality are classed as *chance* variations. Such variations obey the laws of probability as inevitably as the flipping of a coin. Little can be done to change these variations except to revise the control process.

In addition to these chance variations other variations in quality, systematic variations, can be attributed to *assignable causes*, such as differences among equipment, workers, materials, time, etc., and the interrelations of each to the others. Knowledge of the behavior of chance variations is the basis of control-chart analysis. If data vary in a pattern that conforms to the applicable statistical distribution, an assumption is made that no assignable causes exist. These data are said to be *in control*. If the data do not follow the statistical distribution, it is considered that assignable causes are at work and that the process is *out of control*. When the process is *in control*, the distribution of variations for large numbers of items is predictable.

According to Duncan (*1*),[1] ". . . a control chart is a device for describing in concrete terms what a state of statistical control is; second, a device for attaining control; and, third, a device for judging whether control has been attained."

The construction and use of a control chart can be explained simply. Samples of a given size are taken randomly from a process at more or less regular intervals and tested. If no assignable causes exist, these test values will be distributed in a definite pattern. If the distribution is normal, the pattern will assume the shape shown by the normal curve in figure 1. These data from the tests on the samples can be used to construct the control chart shown in figure 2. Actually, the normal curve has been tipped on its side and the values plotted in the horizontal plane on a time or unit of production basis. If a signficiant number of average ($\bar{X}$) values are plotted and all of them fall within the control limits and are randomly distributed about the average, that is, show no trends, then it can be said that the process is in a state of statistical control. If the data do not conform to this pattern, but show cycles or runs above or below average or outside the limit lines, the process should be investigated and the assignable causes ascertained. If all points fall within the control limits and are randomly distributed, assignable causes are not necessarily absent. However, the conclusion that chance causes alone are at work is acceptable, and a search for assignable causes would be unprofitable.

Types of Control Charts

Control charts now in use fall into two categories: Charts of attributes and charts of variables. Usually, control charts of attributes are associated with the *go* or *no go* discipline. They show one of two types of information: (1) The fraction defective or percent of items conforming or nonconforming to specifications, or (2) the number of defective items. Although the attributes may often be measurements,

---

[1] Italic numbers in parentheses identify the references listed on p. 60.

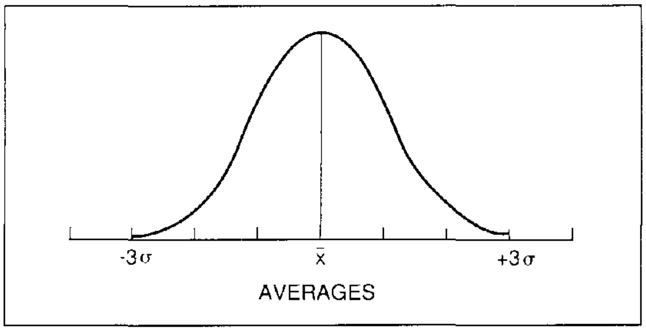


Figure 1.—Distribution of chance variations in a test sample measure of quality.

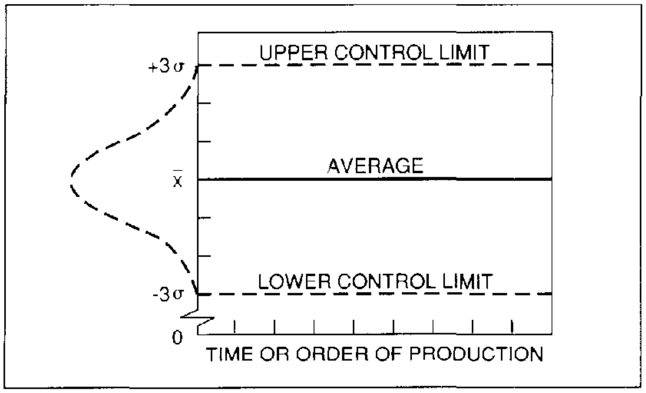


Figure 2.—Theoretical basis for a control chart.

mostly they are visually inspected properties like finish, scratches, missing parts, cracks, or functional ability. On the other hand, control charts of variables are used for values such as p.s.i. density, thickness, length or other such units reported from testing.

Whether to use plans for attributes or variables to control a process in highway construction cannot be answered here. The use of either will be dictated by the process or product to be controlled, and each will provide an effective tool to guide the State and the contractor. The only requisites are that the process be continuous, that the chart be visibly displayed, and that the charts be available to everyone concerned and correctly interpreted. Only control charts for variables will be considered in the remainder of this article.

Control charts can be constructed from test data in the form of individual measurements, averages, standard deviations, and ranges, either singly or in combinations. Again, the form to use will be dictated by the process to be controlled and the economics of the project.

Variables charts fall into two categories: (1) standards given and (2) standards unknown. The usual way to develop a control chart for the standards-given category is to analyze past data and use the information obtained to set up the charts for present and future production. Analysis of the past data will provide values of the average, $\bar{X}$,′ the range R,′ and the standard deviation σ, of the population for calculating the midpoint and limit-line positions on the chart. Usually, the charts are used in pairs—$\bar{X}$ and R or $\bar{X}$ and σ.

Computation of Limit Lines

To determine the respective limit lines for a normal population, when standards are known, the following formulas (*2*) are used:

$\bar{X}$ Chart:

$$\text{Center line} = \text{℄} = \bar{X}'$$
$$\text{Upper control limit} = \text{UCL} = \bar{X}' + A\sigma'_z$$
$$\text{Lower control limit} = \text{LCL} = \bar{X}' - A\sigma'_z$$

R Chart:

$$\text{℄} = d_2\sigma'_x$$
$$\text{UCL} = D_2\sigma'_x$$
$$\text{LCL} = D_1\sigma'_x$$

σ Chart:

$$\text{℄} = C_2\sigma'_x$$
$$\text{UCL} = B_2\sigma'_x$$
$$\text{LCL} = B_1\sigma'_x$$

Where the constants $A, B_1, B_2, D_1, D_2, C_2$, and d_2 for different sample sizes are given in tables of control chart constants like table 1.

Sometimes control charts must be constructed when historical data are not available for analysis and, consequently, the parameters (standards) are unknown. If the standards are unknown, data from 20 to 25 samples, consisting of four to five measurements each, are needed to compute the limits for the charts. Averages, ranges, or standard deviations of measurements on each sample are averaged to give $\bar{X}$, $\bar{R}$ or $\bar{\sigma}$, respectively. Again, the charts are used in pairs—$\bar{X}$ and R or $\bar{X}$ and σ. The following formulas can be used to compute the center lines and the control limits for the charts:

If $\bar{X}$ and R charts are used:

$\bar{X}$ Chart:

$$\text{℄} = \bar{\bar{X}}$$
$$\text{UCL} = \bar{\bar{X}} + A_2\bar{R}$$
$$\text{LCL} = \bar{\bar{X}} - A_2\bar{R}$$

R Chart:

$$\text{℄} = \bar{R}$$
$$\text{UCL} = D_4\bar{R}$$
$$\text{LCL} = D_3\bar{R}$$

If $\bar{X}$ and σ charts are used:

$\bar{X}$ Charts:

$$\text{℄} = \bar{X}$$
$$\text{UCL} = \bar{\bar{X}} + A_1\bar{\sigma}$$
$$\text{LCL} = \bar{\bar{X}} - A_1\bar{\sigma}$$

σ Charts:

$$\text{℄} = \bar{\sigma}$$
$$\text{UCL} = B_1\bar{\sigma}$$
$$\text{LCL} = B_3\bar{\sigma}$$

The values of constants A_1, A_2, B_3, B_4, D_3, and D_4 are listed in table 2.

Limits established by use of the factors from tables 1 and 2 are based on statistical probability and will satisfy the statistician but not necessarily the engineer. Usually the engineer wants the product to meet specification requirements that are based on engineering considerations, and he may not care whether the process is in statistical control. Many engineers prefer that the specification be used to develop the control chart. This can be done, for example, on a two-way specification by using the target as the centerline and the tolerances as the upper and lower limits. If the test results show nonconformance, either the process needs to be changed; sampling and testing techniques altered, processing improved, or the specifications changed to meet existing conditions. The engineer is thus faced with two alternatives: either change the specifications or insist that the requirements be met.

How the Chart Works

The average, $\bar{X}$, chart shows shift of process average. The range, R, or standard deviation σ, charts show process dispersion or spread. Usually, the range, R, chart rather than the standard deviation, σ, chart is used to measure subgroup dispersion. Although both present similar trends, the range values for the R chart are much easier to compute and explain. Charts of individual test values can be used to depict both the

Table 1.—Factors for computing control chart lines—standards given

Number of observations in sample, n	Chart for averages	Chart for standard deviations			Chart for ranges		
	Factors for control limits	Factors for central line	Factors for control limits		Factors for central line	Factors for control limits	
	A	c_2	B_1	B_2	d_2	D_1	D_2
2	2.121	0.5642	0.000	1.843	1.128	0.000	3.686
3	1.732	0.7236	0.000	1.858	1.693	0.000	4.358
4	1.500	0.7979	0.000	1.808	2.059	0.000	4.698
5	1.342	0.8407	0.000	1.756	2.326	0.000	4.918
6	1.225	0.8686	0.026	1.711	2.534	0.000	5.078
7	1.134	0.8882	0.105	1.672	2.704	0.205	5.203
8	1.061	0.9027	0.167	1.638	2.847	0.387	5.307
9	1.000	0.9139	0.219	1.609	2.970	0.546	5.394
10	0.949	0.9227	0.262	1.584	3.078	0.687	5.469
11	0.905	0.9300	0.299	1.561	3.173	0.812	5.534
12	0.866	0.9359	0.331	1.541	3.258	0.924	5.592
13	0.832	0.9410	0.359	1.523	3.336	1.026	5.646
14	0.802	0.9453	0.384	1.507	3.407	1.121	5.693
15	0.775	0.9490	0.406	1.492	3.472	1.207	5.737
16	0.750	0.9523	0.427	1.478	3.532	1.285	5.779
17	0.728	0.9551	0.445	1.465	3.588	1.359	5.817
18	0.707	0.9576	0.461	1.454	3.640	1.426	5.854
19	0.688	0.9599	0.477	1.443	3.689	1.490	5.888
20	0.671	0.9619	0.491	1.433	3.735	1.548	5.922
21	0.655	0.9638	0.504	1.424	3.778	1.606	5.950
22	9.640	0.9655	0.516	1.415	3.819	1.659	5.979
23	0.626	0.9670	0.527	1.407	3.858	1.710	6.006
24	0.612	0.9684	0.538	1.399	3.895	1.759	6.031
25	0.600	0.9696	0.548	1.392	3.931	1.804	6.058
Over 25	$\frac{3}{\sqrt{n}}$	....	(1)	(2)	....	....	....

[1] $1 - \frac{3}{\sqrt{2n}}$.

[2] $1 + \frac{3}{\sqrt{2n}}$.

shift of average and the dispersion, though not as efficiently as the combination of average and range charts.

Following are a few of the many warning signals that have developed to indicate out-of-control processes (*3*):

Individual charts

- 1 test value greater than 2.33 standard deviation from centerline.
- 3 consecutive test values greater than one standard deviation above or below the centerline.
- 11 consecutive test values on same side of centerline job mix target.

Average and range charts

- 1 average or range value outside control limit lines on R and $\bar{X}$ chart.
- 2 consecutive averages outside 2 σ limits above or below the centerline.
- 7 consecutive values on either side of centerline for $\bar{X}$ charts or above centerline for R charts.

These signals are based on statistical probabilities; however, a 3- or 4-point consecutive trend toward either limit or above or below the centerline should be considered a warning. Diagnosis of the reason for the process being out of statistical control is an engineering consideration. The chart can only point out that something is wrong.

Examples of Use

At present, several States are experimentally using variables control charts on several items of construction.

In 1967, the South Carolina State Highway Department (*4*) used specification-based control charts to control and accept production of bituminous hot mix on an Interstate job. A chart on which extraction test results were plotted as individual measurements, averages of 5 per lot, and range of 5 per lot, was used. Charts were also prepared for bitumen content and each of the following sieves: ¾ in., ½ in., ⅜ in., No. 4, No. 8, No. 30, No. 100, and No. 200 for each 2,000-ton lot. A facsimile of a control chart used on the South Carolina project for controlling percent passing the No. 30 sieve is shown in figure 3. Similar charts were used both for the remaining sieves and for bitumen content. Control limit lines for individual measurements, averages, and ranges were superim-

Table 2.—Factors for computing control chart lines—standards unknown

Number of observations in sample, n	Chart for averages		Chart for standard deviations			Chart for ranges		
	Factors for control limits		Factor for central line	Factors for control limits		Factor for central line	Factors for control limits	
	A_1	A_2	c_2	B_3	B_4	d_2	D_3	D_4
2	3.760	1.880	0.5642	0.000	3.267	1.128	0.000	3.267
3	2.394	1.023	0.7236	0.000	2.568	1.693	0.000	2.575
4	1.880	0.729	0.7979	0.000	2.266	2.059	0.000	2.282
5	1.596	0.577	0.8407	0.000	2.089	2.326	0.000	2.115
6	1.410	0.483	0.8686	0.030	1.970	2.534	0.000	2.004
7	1.277	0.419	0.8882	0.118	1.882	2.704	0.076	1.924
8	1.175	0.373	0.9027	0.185	1.815	2.847	0.136	1.864
9	1.094	0.337	0.9139	0.239	1.761	2.970	0.184	1.816
10	1.028	0.308	0.9227	0.284	1.716	3.078	0.223	1.777
11	0.973	0.285	0.9300	0.321	1.679	3.173	0.256	1.744
12	0.925	0.266	0.9359	0.354	1.646	3.258	0.284	1.716
13	0.884	0.249	0.9410	0.382	1.618	3.336	0.308	1.692
14	0.848	0.235	0.9453	0.406	1.594	3.407	0.329	1.671
15	0.816	0.223	0.9490	0.428	1.572	3.472	0.348	1.652
16	0.788	0.212	0.9523	0.448	1.522	3.532	0.364	1.636
17	0.762	0.203	0.9551	0.466	1.534	3.588	0.379	1.621
18	0.738	0.194	0.9576	0.482	1.518	3.640	0.392	1.608
19	0.717	0.187	0.9599	0.497	1.503	3.689	0.404	1.596
20	0.697	0.180	0.9619	0.510	1.490	3.735	0.414	1.586
21	0.679	0.173	0.9638	0.523	1.477	3.778	0.425	1.575
22	0.662	0.167	0.9655	0.534	1.466	3.819	0.434	1.566
23	0.647	0.162	0.9670	0.545	1.455	3.858	0.443	1.557
24	0.632	0.157	0.9684	0.555	1.445	3.895	0.452	1.548
25	0.619	0.153	0.9696	0.565	1.435	3.931	0.459	1.541
Over 25	$\frac{3}{\sqrt{n}}$	....	....	(1)	(2)	....	....	....

[1] $1-\frac{3}{\sqrt{2n}}$.

[2] $1+\frac{3}{\sqrt{2n}}$.

posed on each chart. Limits for average and range were computed from standards known, using appropriate values from table 1 multiplied by a factor of (2.33/3.00) times the standard deviation to reduce them to 98 percent probability intervals. Limits on the chart for the No. 30 sieve were developed from a standard deviation of 2.5 percent, which had been determined in previous research. The limits were computed as follows:

Individual measurements at 98-percent level

$$\text{Job mix tolerance} = (2.33) \times (2.5) = 5.8$$

$$\therefore \quad \text{UCL} = 38.0 + 5.8 \text{ or } 43.8 \text{ percent}$$

$$\text{LCL} = 38.0 - 5.8 \text{ or } 32.3 \text{ percent}$$

Averages of 5 per lot

$$\text{Job mix tolerance} = \left(\frac{2.33}{3.00}\right) \times (A) \times (2.5)$$

$$= (0.78) \times (1.34)$$

$$\times (2.5) = 2.6$$

$$\therefore \quad \text{UCL} = 38.0 + 2.6 \text{ or } 40.6 \text{ percent}$$

$$\text{LCL} = 38.0 - 2.6 \text{ or } 35.4 \text{ percent.}$$

Control limits for ranges were not established on this particular project. However, the central value and upper limit for 5 per lot would be computed as follows for the 99 percent level:

Range of 5 per lot

$$\text{Central value} = (d_2) \times (\text{standard deviation}) \times (\text{reduction factor})$$

$$= (2.326) \times (2.5) \times \left(\frac{2.33}{3.00}\right)$$

$$= 4.5 \text{ percent}$$

$$\text{Upper control limit (UCL)} = (D_2) \times (\text{standard deviation}) \times (\text{reduction factor})$$

$$= (4.918) \times (2.5)\frac{2.33}{3.00}$$

$$= 9.5 \text{ percent}$$

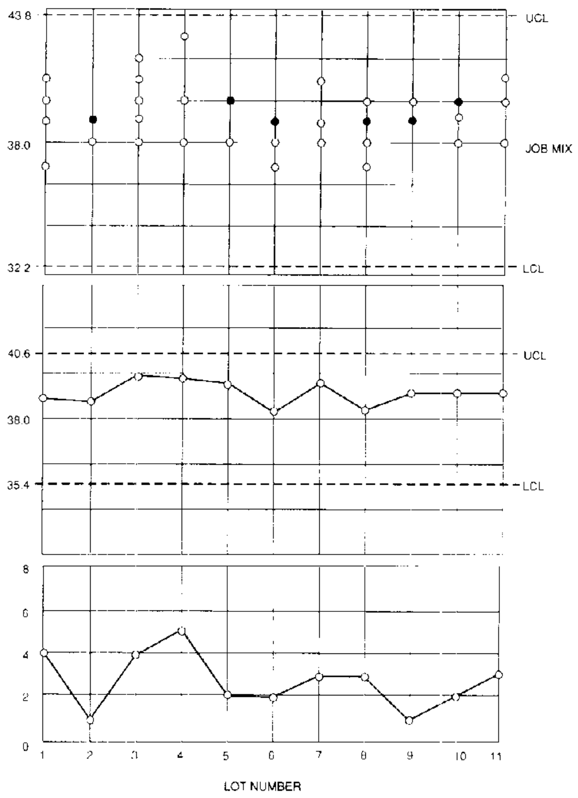


Figure 3.—Control chart used by South Carolina.

Statistically, the control chart for averages (fig. 3) was out of control as seven consecutive points were above the job mix target. However, on this job nothing was done; the situation probably was one of engineering judgment.

Several other States have or are proposing to use control charts. The Vermont Department of Highways, for example, plots all field test data on charts. The California Division of Highways is proposing to use a moving average of five as the basis for corrective action. Trends are more apparent with the moving average chart than with the average chart. Decisions on conformance of the sublot are based on individual-tolerance requirements for the test and on reduced-tolerance requirements for averages of the preceding four sublots plus the one under consideration. On the other hand, the Virginia Department of Highways proposes to plot only individual results and rely on individual probability trends for judgment of conformance. Initially, this approach is being applied experimentally to asphalt mixes. Kansas also plots individual bituminous extraction data and is starting to plot averages of the same data.

For the past 5 years, the Mississippi State Highway Department, the forerunner in the use of control charts, has been using a progressive-step chart to accept and control density of stabilized bases. The 1967 specification extends this use to other bid items. This seemingly complicated plan allows acceptance to be based on as few as two tests per lot, but a minimum of six tests are usually required for rejection. The plan has a built-in safeguard for excluding wild test results. In addition, the test results of the lot are averaged with all test results from the previous five lots, and this average must not deviate by more than a specified amount from the target value.

Summary

The use of control charts to regulate and accept highway construction is still in a fledgling stage. Up to now the industry has existed and expanded with the single test, retest, and engineering-judgment concept. The use of control charts will not alter the fact that the engineer has the final decision in accepting construction. The charts, however, are effective tools to visually forewarn that undesireable trends may be developing and to help both the engineer and the contractor decide *when to take action* and *when not to take action.*

The adoption of control-chart use by the highway industry would provide a significant technological advance in ascertaining adequacy of construction. Almost anyone can spot large changes, but successive small changes that develop as trends are not easily detected, even by experienced people. The picture portrayed by the control charts would highlight the trends and motivate the use of such charts as a decision tool.

REFERENCES

(1) Quality Control and Industrial Statistics, by Acheson J. Duncan, Ph. D., Richard D. Irvin, Inc., 1959, p. 316.

(2) Quality Assurance Through Process Control and Acceptance Sampling, U.S. Department of Transportation, Federal Highway Administration, Bureau of Public Roads, April 1967.

(3) Examples of Control Charts, by C. S. Hughes, Virginia Highway Research Council, Charlottesville, Va., April 1968.

(4) Control and Acceptance of Hot Mix Asphalt Pavements by Statistical Methods in South Carolina, Oren S. Fletcher, February 1968 (to be published in Highway Research Record No. 290, HRB, November 1969).

(5) ASTM Manual on Quality Control of Materials, STP-15–C, American Society for Testing Materials, January 1951.

☆U.S. Government Printing Office : 1990 - 517-780/20898