Pavement Performance as Affected by Materials Testing Program

THOMAS W. KENNEDY, HAROLD L. VON QUINTUS, and TERRY MITCHELL

ABSTRACT

A summary is presented of a study aimed at providing a means by which an organization can determine the cost-effectiveness of individual tests, test methods, and associated sampling frequencies for tests used in controlling the quality of pavement materials as related to effects on performance. Appropriate procedures have been developed and are briefly discussed, including critical considerations and limitations due to lack of suitable stochastic models to predict performance and contractor response to changes in testing frequency. These procedures are embodied in the computer program COSTOP1, which was developed to assist state highway agencies in determining the optimum test frequency for a single test or the optimum test program for multiple tests to produce the greatest return for every dollar spent on testing. COSTOP1 is general and modular so that testing programs for all paving construction and materials can be evaluated, and new models and differing repair strategies may be easily defined and input. Preliminary results indicate that higher frequencies of testing than commonly used would be cost-effective, decreasing the equivalent annual pavement costs by much more than the additional testing costs.

Quality assurance testing programs practiced in most states are labor intensive and costly. These programs are estimated to cost approximately $200 million per year for federal-aid highway projects alone. Unfortunately, it is not certain that these programs guarantee quality or that better quality could not be attained.

Material and construction specifications often were developed by a "trial and error" process supplemented by research and development efforts and were aimed at duplicating successful projects. Tests were standardized and experience with their use and consequent results were shared through direct contact, technical meetings, and technical publications. This evolutionary process produced a variety of quality control tests and testing frequencies.

The types and frequencies of tests conducted and the test procedures vary widely among the various highway agencies that use a variety of different materials and design procedures for different environmental conditions and traffic volumes. In addition, the specification values assigned to specific tests and the recommended frequencies of testing often were based on the judgment of individual engineers because it is difficult to objectively relate control test values and the frequency of testing to actual pavement performance. This is due in part to the complexity of pavement behavior and analyses, which is caused by numerous interacting factors such as construction procedures, environmental factors, traffic differences, and variable materials.

Although there has always been a need for quality assurance programs, escalating costs, increased traffic loads and volumes, and recent administrative and legislative pressures to keep costs and personnel levels at a minimum have resulted in research aimed at improved specifications and more cost-effective quality assurance programs. The Federal Highway Administration established Project 6-G, "Performance Related Specifications for Highway Construction and/or Rehabilitation" as a coordinated effort to accomplish these goals. Accordingly, a study was initiated to provide a basis or methodology to determine and evaluate the cost-effectiveness of individual tests and associated sampling frequencies used in controlling the quality of pavement materials as related to pavement performance. A summary of this study (1) is the subject of this paper.

STUDY OBJECTIVE

The primary objective of the study summarized in this paper was to provide a means or methodology to determine the cost-effectiveness of individual tests, test methods, and associated sampling frequencies used in controlling the quality of pavement materials as related to performance. The resulting methodology provides a means for individual agencies to answer the following questions:

1. Does a specific test or group of tests provide information that directly relates to performance and how sensitive is performance to variations in test values that could normally be expected to occur?

Certain tests may not measure properties that are important to the performance of the pavement, or the performance may not be significantly affected by variations that might occur. Thus, conducting the test may not be cost-effective, or the expected variation or the effect of such variation may be so small that additional testing is not cost-effective after a given level has been established.

2. How frequently should tests that relate to performance be conducted?

Most specifications concerning test frequency are established with respect to statistical concepts related to obtaining an accurate estimate of the engineering properties being measured. Although this is important, a more important question relates to how frequency affects the quality of materials and construction, which in turn will affect the performance of the pavement. Thus, will additional tests cause a contractor or supplier to provide a higher quality product? In general, it is assumed that more tests will cause an improvement in quality, but at best there is a point of diminishing returns that is also a function of how the specifications are written and of the penalties associated with poor-quality materials and construction.

METHODOLOGY

To evaluate the cost-effectiveness of alternative tests or testing plans, the probable benefits of each test frequency and the costs of the particular test being evaluated must be determinable. The relative benefits will accrue from increased performance or lower maintenance and rehabilitation costs, and the relative cost will be the increased cost either of performing more tests or of changing to a test procedure that has a better relationship to performance. Of course, just the opposite situation could also be cost-effective. It may be possible to reduce testing frequency and testing cost while only marginally decreasing performance with an attendant slight increase in maintenance and rehabilitation costs.

To determine the cost-effectiveness of a particular test frequency, five primary questions must be answered:

1. What effect does the material property (both mean and standard deviation) have on pavement performance?
2. What is an acceptable variation of the material property?
3. Will construction practices or production techniques be affected by altering the testing and sampling frequency and, if so, what is the effect?
4. What is the cost of testing to measure the material property?
5. What is the cost to maintain, repair, or replace defective material that was accepted?

The resulting methodology, which is a computer program entitled COSTOP1, considers the relationship between the test value and the performance of the highway, the effect of variations from the specifications, and the consequences of accepting unsuitable material. Thus state DOTs can establish priorities among quality control tests and optimize sampling frequencies for each test. This includes the necessary level a test parameter must achieve to provide the desired quality and the frequency at which testing must be performed to minimize the possibility of having defective materials because of poor construction practice.

The methodology is applicable to all highway-related materials and soils; however, the study was limited to asphalt mixtures and concentrated on those properties that are used as standards by most state agencies.

The critical topics considered are

1. Models relating quality control test results to pavement performance;
2. Variation in materials properties, which includes (a) the effect of testing frequency on knowledge of the pavement and on contractor performance and (b) the relationship between frequency of testing and anticipated consequent pavement performance;
3. Cost of materials testing;
4. Cost associated with repairing defective materials and pavements; and
5. Comparison of the differential cost of testing to the differential pavement costs by means of benefit-cost analyses.

Each of these topics was carefully studied and evaluated with respect to how each was to be handled in the methodology. Figure 1 is a simplified flow chart illustrating the cost-effectiveness methodology.

Pavement Performance and Predictive Models

The concept of basing a materials test program and its optimization directly on its effects on the performance of the pavement is logical, appropriate, and desirable. Unfortunately, the mathematical models needed to support the methodology do not exist.

Review of the literature produced only a few models that predicted distress on the basis of material properties derived from conventional testing, but these were quite limited. Models were more plentiful that predicted pavement distress or performance with relation to material properties obtained from more sophisticated test procedures not commonly conducted, such as resilient moduli, fatigue potential, and permanent deformation potential. There were also a few models that predicted stiffness or resilient modulus in terms of conventional test results.

Because of the shortage of models to predict pavement distress or performance in terms of material properties measured by standard quality control or acceptance tests, it was necessary to couple two models, one to predict an engineering property in terms of measured material properties and the other to predict distress or performance. The models selected for use include

1. The asphalt concrete mixture stiffness model developed by Miller et al. (_2_). This model is used to predict stiffness as an independent variable for other models.
2. The Arizona model for roughness (_3_), which is based indirectly on asphalt content and gradation of the aggregate in the asphalt concrete mixture.
3. The Waterloo models (_4_) for rutting and for strain at the bottom of the asphalt concrete as functions of asphalt concrete stiffness, pavement structure, stiffness of the subgrade, and traffic.
4. Algorithms for predicting present serviceability index and fatigue cracking from the mechanistic model VESYS IV-B (_5_).
5. A model predicting loss of skid resistance developed by Roberts and Jordahl (_6_).

The approach taken was twofold: (a) the computer program embodying the methodology was to be modular so that it could accept any model and (b) limited models were to be developed as part of the project with the sole intent of demonstrating the methodology. The approach used to develop models was to combine two models, one of which would predict an engineering property such as material stiffness in terms of conventional test results and the other of which would predict a particular distress in terms of the engineering property so derived.

The simple approach described produces a deterministic model without the capability of considering variability in the material properties that could be affected by the test program. The deterministic equations were transformed into stochastic equations by expansion into first-order Taylor series, allowing the propagation of variance. Resulting stochastic equations for predicting pavement distress or performance in terms of material properties and

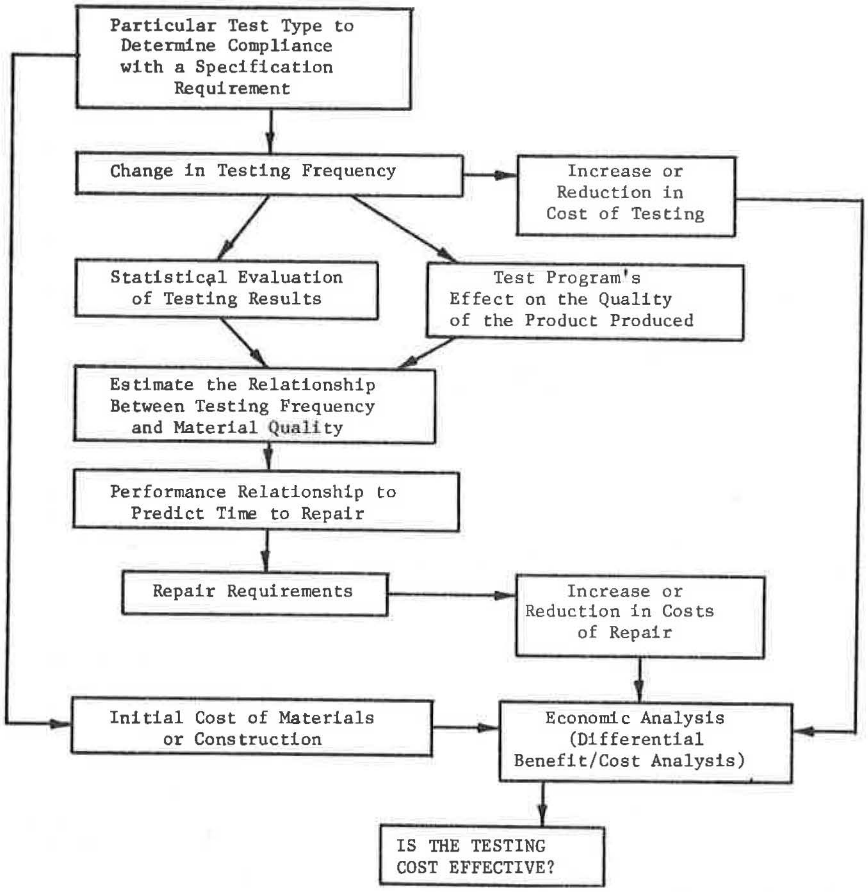

FIGURE 1 Simplified flow chart illustrating cost-effectiveness methodology to determine optimum test frequency for a particular test.

their variations from conventional testing are believed adequate for demonstrating the methodology, but no claim is made for their adequacy for general use.

General Variability in Pavement Materials

It is a generally recognized statistical rule that the accuracy with which the mean value of a population may be estimated increases with the number of samples from the population measured. The accuracy of the estimate for standard deviation or variability from the mean also increases with sample size. It follows then that the greater the number of material tests conducted, the higher the confidence level that the mean will be identified with sufficient accuracy, that the variability will be better defined, and that substandard materials will be identified. This logic leads to the question of how many tests should be conducted to satisfactorily identify the characteristics of the material. This question implies some consideration of the significance of the test as well as the number for statistical sufficiency. These subjects were dealt with in the study.

There is an assumption necessary for optimization of number of tests, and that is that ability to better evaluate material properties is rewarded by improvement in the properties themselves. Here again there were lack of information and lack of models to relate the effects of number of tests in a test program to actual material properties produced (as opposed to more accurate evaluation of the mean and variance). Therefore this assumption is necessary but cannot be corroborated without long-term data collection to verify it.

Variability Resulting from Contractor Responses to Testing Frequency

Another factor that tends to confound efforts to optimize sampling and testing frequency is the effect on a contractor of testing frequency. Superficially, it appears simple that a contractor would be expected to produce a superior product with the knowledge that testing frequency is high. Superior in this sense is defined not only as always exceeding specified minimum test values but also as maintaining a reasonable level of uniformity. This factor is conceptually shown in Figure 2, Curve A.

Discussion of this concept with state highway agency officials revealed that some officials believe testing frequency has no effect on contractor performance (Figure 2, Curve B). Others believe that, if data were assimilated and plotted, the resulting curve would show that material variability would appear to increase with testing frequency. This phenomenon is believed to occur because state agencies tend to subject contractors with poor control histories to greater amounts of testing. The result, shown conceptually in Figure 2, Curve C, is that apparent variability of paving materials may statistically increase with testing frequency due to the correlation of two related effects and not as a cause and effect relationship.

Another complexity is the presence of lag time in

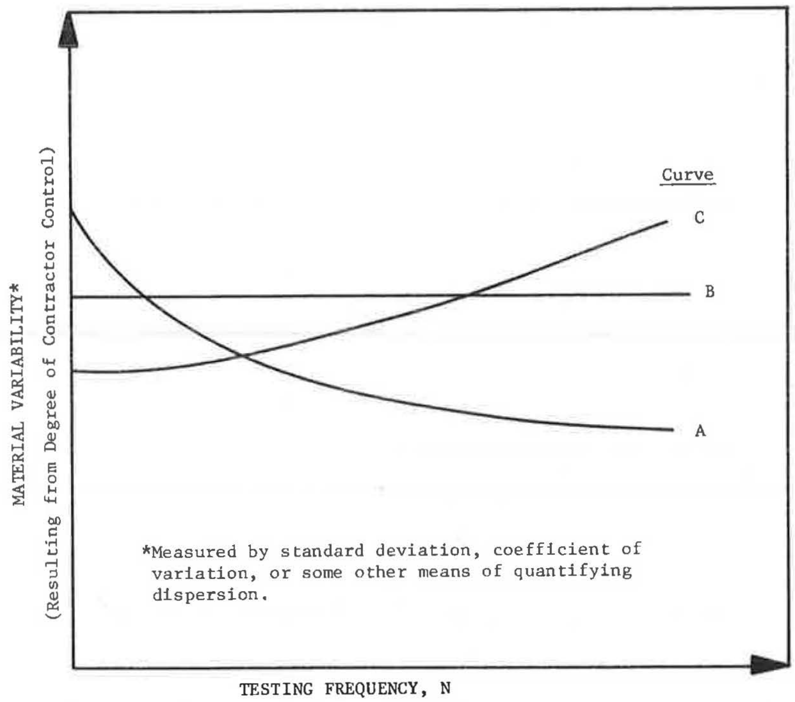

FIGURE 2 Effect of test frequency on material variability.

contractor response. For instance, it is doubtful that the effects of a change in the established testing program for a state DOT or individual highway district would appear immediately on current projects. It is more likely that these effects would appear over a period of time on later projects and in varying fashions for different contractors.

Unfortunately, the complexities of the responses of a single contractor or contractors in general are not well understood and have not been committed to mathematical models that could contribute to this project. These responses are likely to include changes in construction costs as well.

The matter of predicting contractor responses in terms of materials properties was evaluated, resulting in the adoption of a general model that would allow inclusion of this consideration in the methodology but leaving broad flexibility for individual users to define this response on the basis of their own experience or perceptions of its nature.

Testing Costs

The effectiveness of the proposed methodology in optimizing materials test programs is dependent on the accurate evaluation of costs per test. It is rather common for state DOTs in estimating costs for their activities to leave out significant indirect costs and even direct costs. Therefore a standard and reasonable methodology was selected through modification of a procedure used by the Louisiana Department of Transportation and Development to ensure meaningful estimates of testing costs. This procedure for determining the costs of quality control and acceptance tests is described in detail elsewhere (1) and includes on a per test basis the following:

1. Salary costs,
2. Equivalent depreciation cost,
3. Vehicle and equipment rental cost,
4. Travel cost,
5. Supply cost,
6. Administrative overhead costs,
7. Administrative engineering costs, and
8. Total cost per test (the sum of the items listed).

Pavement Rehabilitation and Maintenance

Assuming that stochastic models are available to predict distresses in terms of variables, including both the means and the variances of material properties, levels of distresses of various types can then be predicted in terms of numbers and types of quality control and acceptance tests. This leads to the necessity of identifying the consequences of distress or deterioration in terms of type of rehabilitation or maintenance that will be required and what it will cost. Because rehabilitation or maintenance strategies and their costs vary widely from state to state and from district to district within a state, it was necessary to develop a flexible system for defining these strategies in terms of levels of distress of various types and for assigning costs for these strategies. The details of this are given elsewhere (1).

With the development of this system, there were available the components necessary for considering (a) initial cost of construction, (b) costs for various materials test programs, (c) effects of materials test programs on performance of pavement, (d) rehabilitation or maintenance strategies that may result from various distress levels, and (e) costs for those rehabilitation or maintenance strategies. These provided the opportunity to optimize costs taking all these factors into consideration.

Cost Analysis

An incremental benefit-cost analysis is used to compare alternate testing schemes. Each alternative has the following costs with which it is associated:

- Testing,
- Construction,

- Maintenance,
- User, and
- Rehabilitation.

For each alternative, construction, maintenance, user, and rehabilitation costs are combined into an equivalent uniform annual cost over the life of the pavement. Similarly, testing costs are converted to equivalent uniform annual costs and alternatives are then arrayed in order of increasing (annual) testing costs. Incremental benefit-cost analysis is conducted using a challenger-defender approach to determine the most economical alternative.

PROGRAM COSTOP1

COSTOP1 is a computer program that simulates the appearance and growth of pavement distress with age and number of vehicles, determines a time at which one or more failure criteria are exceeded, and evaluates the economic consequences of such functional failures. The calculated results are tied to a testing program for materials used in construction by varying the values, standard deviations, or both, for test results as functions of the number of tests performed. The program can analyze a large number of testing programs and perform a differential benefit-cost analysis on the results to indicate to the user the most beneficial test programs, subject to the assumptions and distress models used.

COSTOP1 calculates equivalent annual pavement and testing costs for each test frequency or test program. All costs are based on a unit of dollars per lane mile. In addition, for each test program the age at failure (the time at which one of the distress values has exceeded the critical value established by the agency), the type of failure (the specific distress type requiring maintenance or rehabilitation), and the type of repair technique selected from the decision tree are all determined for the user.

The program also prints the differential benefit-to-cost ratio for each testing alternative considered. The benefit is the decrease in equivalent annual pavement costs because of an increase in the testing program, and the cost is the increase in testing costs between two test programs. All possible combinations of testing frequencies are not printed; only those alternatives with a differential benefit-to-cost ratio greater than 1 are printed.

DEMONSTRATION OF EVALUATION OF SELECTED TESTS

Using data accumulated from one state, computer program COSTOP1 was used to determine the cost-effectiveness of selected asphalt concrete tests. These tests include

- Mix gradation (percentage passing No. 200 sieve),
- Asphalt concrete thickness,
- Percentage air voids (or compaction),
- Asphalt content,
- Asphalt viscosity (or penetration), and
- Los Angeles abrasion.

In most cases more than one quality characteristic or test value must be considered in defining the most optimum test program. For example, asphalt concrete thickness alone is insufficient to assure the desired performance. To be durable, an asphalt concrete mixture must also have the necessary amount of bitumen and proper grading.

An evaluation of all tests performed in the central, district, and residency asphalt laboratories was impossible because performance functions relating each test to pavement performance (either directly or indirectly) were unavailable. However, the techniques presented can be used for any test or combination of tests or with any other construction material provided performance models and testing cost data are available or can be determined.

In summary for this example, Table 1 gives the optimum number of tests for each test listed. As shown, bitumen content and asphalt concrete thickness are the two most important tests for this example. To demonstrate the difference between rural and urban areas (different traffic levels and corresponding pavement cross sections), data from one state were used to compare the optimum testing program for different highway types (state routes, U.S. routes, and Interstate highways). A summary of these results is given in Table 2. As shown, the optimum testing program varies depending on the highway type. For the particular example evaluated, surface thickness is the most cost-effective test for low-volume roads, and percentage air voids is the most cost-effective for high-volume roadways. In addition to these two examples, the decision criteria used by a state agency to manage its pavements will also have an effect on the selection of optimum test programs. Table 3 gives a summary of the results for one state that uses rut depth as its primary decision criterion. As shown, if the critical rut depth that causes maintenance is changed from greater than or equal to 0.5 in. for 50 percent of the wheelpath area to greater than or equal to 0.75 in. for only 25 percent of the wheelpath area, the least cost-effective test changes from asphalt viscosity to gradation.

TABLE 1 Summary of the Optimum Testing Program for an Asphalt Laboratory

Test	No. of Tests per Lane Mile (asphalt concrete tonnage = 1,750)
Bitumen content	9
Percentage air voids	6
Percentage passing No. 200 sieve	3
Asphalt viscosity	3
Asphalt concrete thickness	9
Los Angeles abrasion	0

TABLE 2 Summary of the Effect of Highway Type on Optimum Test Programs (asphalt concrete tonnage = 1,750)

Test	Low Traffic	Moderate Traffic	High Traffic
Extractions/percentage air voids (density)	3	6	12
Asphalt viscosity	3	3	9
Asphalt concrete thickness (cores)	6	9	3

CONCLUSIONS

The primary objective of the study was to develop a capability for determining the cost-effectiveness of individual tests and associated sampling frequencies used in controlling the quality of pavement construction and materials as related to performance. This objective was generally accomplished, but immediate implementation is limited because of the lack

TABLE 3 Summary of the Effect of Critical Distress Criteria on Optimum Test Program

Type of Test	Rut Depth > 0.50 in. for 50% of Wheelpath (HMAC t = 4.0 in.)	Rut Depth > 0.75 in. for 25% of Wheelpath for HMAC Thickness (in.) 4.0	1.5
Bitumen content	10	10	3
Percentage air voids	10	10	3
Percentage passing No. 200 sieve	10	3	3
Asphalt viscosity	3	10	3
Asphalt concrete thickness	10	10	10
Failure mode	Rut depth	PSI	Damage index
Asphalt concrete tonnage per lane mile	1,600	1,600	500

of models relating materials properties commonly measured to performance of a pavement or other product.

In addition, a serious conceptual problem was encountered in discriminating between the statistical concept of improving estimates of means and standard deviations compared to actual (or "population") means and standard deviations and the actual effect of increased testing on a subsequent product to be produced later. There is no doubt that an increase in testing frequency offers a better opportunity for identifying and perhaps replacing deficient materials. It also appears logical that the contractor will respond by producing a better product on subsequent portions of the current project or on other projects for which he expects high test frequencies. However, there are certainly no established relationships that indicate what a typical contractor response would be, let alone what a specific contractor might do under a specific set of conditions.

According to the limited models available for demonstration, the methodology appears to consistently indicate that high testing frequencies are cost-effective. This appears logical (almost obvious) in view of the relatively nominal cost of testing compared to costs for repair and rehabilitation. It generally requires little improvement in the product, especially in reduction of variance, to increase service life before required repair or rehabilitation.

ACKNOWLEDGMENT

The study on which this paper is based was supported by the Federal Highway Administration under Contract DTFH61-82-C-00015.

REFERENCES

1. H.L. Von Quintus, J.B. Rauhut, T.W. Kennedy, and P.R. Jordahl. Cost Effectiveness of Current Sampling and Testing Programs for Paving Construction and Materials. Draft Final Report. FHWA, U.S. Department of Transportation, Aug. 1984.
2. J.S. Miller, J.Uzan, and M.W. Witzcak. Modification of the Asphalt Institute Bituminous Mix Modulus Predictive Equation. In Transportation Research Record 911, TRB, National Research Council, Washington, D.C., 1983, pp. 27-36.
3. G.B. Way and H. Jones. Asphalt Concrete Variance and What It Means in Terms of Future Pavement Performance. Report 1979-GW2. Arizona Department of Transportation, Tucson, Aug. 1979.
4. F.R.P. Meyer and R.C.G. Haas. A Coordinated Method for Structural Distress Predictions in Asphalt Pavements. Presented at the Annual Meeting of the Association of Asphalt Paving Technologists, Feb. 1978.
5. P.R. Jordahl and J.B. Rauhut. Flexible Pavement Model VESYS IV-B. Draft Final Report. FHWA, U.S. Department of Transportation, Aug. 1983.
6. F.L. Roberts and P.R. Jordahl. Use of Flexible Pavement Skid Resistance Measurements in Developing Pavement Damage Functions for Cost Allocation. Brent Rauhut Engineering, Inc., Austin, Tex., Jan. 1983.

Publication of this paper sponsored by Committee on Quality Control and Acceptance Procedures.