Evaluation of California Profilograph

LARRY A. SCOFIELD, SYLVESTER A. KALEVELA, AND MARY R. ANDERSON

The Arizona Department of Transportation evaluates portland cement concrete pavements by testing with mechanical as well as electronic profilographs. The precision of the two types of profilograph was evaluated. More than 100 profilograph runs were conducted on a selected pavement section. The range of replicate readings of pavement profile index could be as much as 2.0 in./mi for a rough pavement. Electronic profilographs adjusted to operate at low filter settings gave lower profile index values than those obtained with the same profilographs at higher filter settings.

The first California profilograph was developed in 1940. During almost 50 years of use, it has seen many changes. The beam length has varied from 10 to 25 ft. It has been a mobile unit and a hand-propelled unit. There have been as many as 16 wheels and as few as 4. It has been constructed of wood, steel, and aluminum, and it has been assembled in three to five sections. The model most prevalent in the industry today resembles the 1962 12-wheel profilograph. During the mid-1980s the recording device was computerized by Cox and Sons, Inc. Both mechanical and computerized versions (automated) are currently available in the industry.

It is reported that California developed and published a 7-in./mi profile index specification between 1958 and 1960 (*1*). The original specification and test procedures are still widely used today. The specification appears to have been established on a limited number of pavement sections built with fixed-form construction, from profiles obtained on the outer wheelpath, in the direction of traffic, with a mobile profilograph.

PROBLEM STATEMENT AND STUDY OBJECTIVES

The California Department of Highways developed the profilograph test equipment and roughness specifications to provide an objective method for ensuring a minimum for ride quality for concrete pavements. These devices and methods, developed more than 30 years ago, were based on subjective ride rating surveys and prepared for convenient and expedient application in the construction environment.

Today, ride-quality specifications have been extended far beyond the intent of the original procedures and specifications. In the past, the 7-in./mi roadway simply represented the minimum ride quality needed. Incentives and disincentives were not used. Incentives today can reach 5 percent of the bid item unit price.

L. A. Scofield, S. A. Kalevela, Arizona Transportation Research Center, Arizona Department of Transportation, 206 South 17th Avenue, Mail Drop 075R, Phoenix, Ariz. 85007. M. R. Anderson, Industrial and Management Systems Engineering Department, College of Engineering and Applied Sciences, Arizona State University, Tempe, Ariz. 85287.

Although the California profilograph and roughness specifications have served the industry well, two questions must be resolved. First, are the test procedure and equipment sufficiently accurate and reproducible to warrant such high percent incentive/disincentive specifications? and second, are the mechanical and computerized profilograph units comparable? Because of these questions, a study was conducted to

- Evaluate the precision of the California profilograph;
- Compare mechanical and automated profilographs;
- Evaluate the effects of data filter settings on profile index readings; and
- Evaluate the effects of trace reading variability on the profile index obtained with the mechanical profilograph.

EXPERIMENTAL DESIGN

To accomplish these objectives three test plans were developed: one to develop a precision statement for profilograph testing, one for analyzing the effect of the data filter settings, and one to determine operator variability in trace reduction.

The main experiment was conducted by using a 4-×-4-×-2 randomized block design with replication, resulting in 64 profilograph runs. The experimental design consisted of four operators, four profilographs (two mechanical and two automated), and two levels of pavement roughness (2.5 and 10 in./mi).

The data filter setting experiment consisted of a 3-×-2-×-2-×-2 randomized block design with replication. The experimental design consisted of three data filter settings (8,000, 6,000, 4,000), two automated devices, two operators, and two roughness levels (2.5 and 10 in./mi). The total number of tests for this experiment was 48.

The trace reduction experiment consisted of a 4-×-2-×-2 randomized block design with replication. The experimental design consisted of four operators, two mechanical devices, and two levels of roughness (2.5 and 10 in./mi). Each operator analyzed eight traces during the first round of reading, and a copy of the same eight traces in the second round.

Test Site Location

The field testing was conducted in Phoenix, Arizona, between September and November 1990. Testing was performed on an undoweled plain jointed concrete pavement 12 in. thick. The concrete pavement was constructed on an aggregate base and used skewed random joint spacing of 13, 15, 17, and 15 ft. The pavement had been constructed approximately 5 months earlier and was not open to the public.

Initially, three levels of roughness were desired to better represent the range of expected roughness levels obtained during construction operations. However, the large number of tests needed for this experiment and the difficulty in finding test sections in proximity for convenient testing resulted in the selection of only two levels of roughness.

Equipment and Operators

Four profilographs were used in this study. Two of the profilographs were mechanical and two were automated (computerized). The automated devices were Cox and Sons, Inc., Model CS8200s. One was a retrofitted McCracken unit and one was an original CS8200 unit. Of the two mechanical units, one was made by Cox and Sons, Inc., and the other was a McCracken device. The oldest of the four profilographs was purchased in 1967, and the newest was purchased in 1989. For this study, the two mechanical devices were identified as M1 and M2 and the two computerized devices were identified as E3 and E4.

The operators used for these experiments represented actual construction operators. Each operator represented a different construction group. Therefore, any real differences in methods or procedures between the groups should be revealed in the variability.

The four operators used in the field testing were not the same as those used for analyzing trace reduction. Problems with personnel availability precluded consistent use of all operators between these two segments of the experiment.

Test Procedures

Precision Experiment (Main Experiment)

The intent of the main experiment was to determine the "actual" field variability as opposed to the "ideal" variability possible with the devices. Therefore, the operators were not instructed on how to conduct the testing; they were only instructed on the run sequences and the manner in which the operators would switch devices to provide randomization. Each operator delivered the device managed by his or her construction unit to the test location. The operators assembled their own devices in their normal manner. Before conducting testing, the research group checked each device after assembly for proper calibration. One-in. calibration blocks were used for checking the vertical calibration for the manual units, and all units were gauged against a presurveyed 528-ft distance calibration check. Each of the measurement wheels was visually checked for eccentricity. It took three separate attempts to accomplish the complete experiment with all devices in satisfactory operating condition. Although the final testing was completed in 8 hr, it took several months to arrange the logistics for mobilizing all four units and operators during the three attempts at conducting the experiment.

Before assembling the profilograph units, a K. J. Law 690DNC profilometer was used to conduct the first series of runs over the test sections. Ten runs were made with the profilometer during the day, representing the time span over which the main experiment was conducted. This provided the ability to evaluate any changes in actual pavement roughness with time of day (i.e., thermal curling). Five runs were made with the profilometer for each of the two test sequences. One test sequence was conducted in the morning and one in the afternoon. Because the profilometer measures the profile in each wheelpath simultaneously, only one run was necessary to obtain both wheelpaths.

Profilograph testing was conducted in two replicates and in a complete randomized block design as much as possible within each replicate. Complete randomization was limited by four operators and four machines tested at about the same time. This allowed continuous testing with all operators and devices while ensuring statistical validity. For each operator and machine combination, testing began on the rough wheelpath in the direction of travel. Upon completion of this run, testing was continued with a subsequent run along the smooth wheelpath in the opposite direction of traffic. This "looping" allowed testing to be conducted without the need for deadheading the equipment. During testing no guides were used to ensure proper tracking.

Trace reduction for the mechanical devices was accomplished using only one individual to minimize trace reduction variability. One of the research engineers at the Arizona Transportation Research Center performed all trace reductions to provide evaluations as consistent as possible.

Data Filter Setting Experiment

Upon completion of the main experiment, the data filter experiment was conducted using two operators and two automated profilographs. As in the main experiment, test sequencing was conducted in a randomized complete block design within replicates, subject to the tests conducted in pairs. This allowed continuous testing with both operators and devices. Again only visual alignment control was used.

Trace Reduction Variability Experiment

Profilograph traces produced with the mechanical devices, M1 and M2, during the first attempt at the main experiment were used in the trace reduction variability experiment. Operators were not instructed regarding trace reduction techniques; they used their own established procedures.

Four operators were each given the same set of eight profilograph traces to interpret. The set consisted of four traces for a smooth pavement surface and four traces for a rough pavement surface. For each pavement type, two traces were obtained with the device M1 and two traces with M2. The eight traces were labeled in random order and given to each operator. After the first trace reductions were completed, a new random order of the same traces was sent to the same operators for reduction. Approximately 1 month passed between reductions. At the time of the first reduction, the operators were not advised about the second set of readings.

All copies of the traces were obtained from the same originals by a Xerox 2080 machine. This machine was selected to alleviate the concern that the final traces would be distorted

when compared with the original traces. This also allowed production of clean, unmarked profiles for each reduction. This precluded any bias that might result from eraser marks on the traces during operator interpretation.

Analysis Procedures and Terminology

Four categories of statistical procedures were applied during data analysis: (a) *F*-test for significance of treatment effects, (b) *t*-test for comparison of two means, (c) Duncan's multiple range tests for multiple comparison of several means, and (d) the standard analysis of response repeatability and reproducibility (*2*). The *F*-test was used to test the significance of treatment effects in the analysis of variance. The Student's *t*-test was used in cases that involved the comparison of two means. Where the desire was to make multiple comparisons of several means, Duncan's multiple range test was used.

For this study, repeatability was defined as the closeness of agreement between mutually independent test results obtained from the same wheelpath within the short time intervals by the same operator with the same device. The smaller the range of the test results, the better the repeatability of the test. Reproducibility was defined as the closeness of agreement between mutually independent test results obtained from the same wheelpath by different operators with the same profilograph.

Because the "true" profile index of each wheelpath of the roadway was not known, it was decided to obtain surrogate reference values. The arithmetic average of all test results for a given wheelpath was taken to be the reference value for that wheelpath. An additional evaluation was made by comparing individual readings to the mean value for each wheelpath of the test roadway.

The terms "track" and "road" have occasionally been used to mean the wheelpath over which profilograph tests were conducted.

EXPERIMENTAL RESULTS

Main Study

The test results of the main study are summarized in Table 1. The analysis of variance for the main study showed that (a) roughness readings produced by the four devices were statistically different at the 1 percent significance level, and (b) different operators produced statistically different profile indexes at a significance level of 7 percent.

In general, the two mechanical devices exhibited slightly better repeatability than the automated devices. That is, for a given combination of operator and device, the mechanical devices provided slightly more consistent results. The results in Table 1 show the repeatability range of each device for the smooth and the rough wheelpaths. All devices were repeatable within 2.0 in./mi on the rough wheelpath and within 1.5 in./mi on the smooth wheelpath. The average repeatability range was 0.75 in./mi for the rough wheelpath and 0.56 in./mi for the smooth wheelpath. This repeatability range is an average computed from the test results produced by the four operators for each device.

The closeness of test results to the means for the smooth and rough wheelpaths is depicted in Figures 1 and 2, respectively.

TABLE 1 PAVEMENT ROUGHNESS DATA OBTAINED WITH PROFILOGRAPHS

		Profile Index Readings (in/mile)					
		Smooth track			Rough track		
Profilograph	Operator	Test 1	Test 2	Difference*	Test 1	Test 2	Difference*
M1	1	4.00	3.50	0.50	10.00	8.50	1.50
	2	3.50	3.50	0.00	8.50	7.50	1.00
	3	5.00	4.00	1.00	7.00	7.00	0.00
	4	4.00	4.00	0.00	7.00	8.00	1.00
	MEAN	4.13	3.75	0.38	8.13	7.75	0.88
M2	1	5.00	5.00	0.00	11.00	11.00	0.00
	2	5.00	6.00	1.00	10.00	11.00	1.00
	3	5.50	6.00	0.50	9.00	9.50	0.50
	4	5.50	5.00	0.50	10.50	11.00	0.50
	MEAN	5.25	5.50	0.50	10.13	10.63	0.50
E3	1	5.00	6.00	1.00	8.50	8.00	0.50
	2	5.00	6.50	1.50	7.50	8.50	1.00
	3	6.00	6.50	0.50	8.00	8.50	0.50
	4	5.50	5.50	0.00	7.50	8.00	0.50
	MEAN	5.38	6.13	0.75	7.88	8.25	0.63
E4	1	6.00	6.50	0.50	11.00	11.00	0.00
	2	7.00	6.00	1.00	9.00	8.50	0.50
	3	7.00	6.50	0.50	8.00	10.00	2.00
	4	6.00	5.50	0.50	10.00	8.50	1.50
	MEAN	6.50	6.13	0.63	9.50	9.50	1.00

* Absolute difference between tests 1 and 2

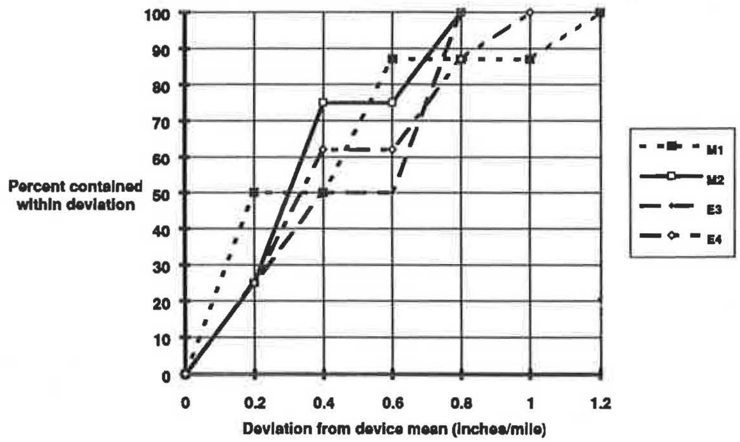

FIGURE 1 Device repeatability depicted in terms of cumulative percentage of device readings within given deviation from device mean on smooth track.

The figures represent, for each device, the percentage of actual test results contained within given deviations from the device mean for the smooth and rough wheelpaths. For example, approximately 100 percent of the test results obtained by all four operators for Device E4 were within a 1-in./mi deviation from the device mean value for the smooth wheelpath. However, only about 60 percent of the test results for Device E4 were within 1 in./mi from the device mean for the rough wheelpath.

The range of roughness values obtained by all operators with each of the devices for the smooth and the rough wheelpath is shown by Figure 3 by solid bars. The mean values for the smooth and the rough wheelpaths are depicted by thick vertical lines. Figure 3 indicates that even though the individual devices may be very repeatable, test results from each of the devices may be significantly different.

The range of test results was between 3.5 and 7.0 in./mi for the smooth wheelpath and between 7.0 and 11.0 in./mi for the rough wheelpath. The quality of these data clearly is not acceptable to administer an incentive/disincentive specification. It should be remembered that the variability could be larger if operator variability in trace reduction were included.

Variability Due to Data Filter Settings

The Cox and Sons Model CS8200 recommends a data filter setting of 8,000. To evaluate the effect of reducing the filter setting, two operators and two automated devices were evaluated at three settings for both the rough and smooth conditions. A total of 48 tests were performed. The results of this testing are given in Table 2. The profile index values for each level of filter setting, given as percentages of the values at the 8,000 filter setting level, are plotted in Figure 4. The values represented in Figure 4 constitute the average of all values obtained at a given filter setting for each track condition. Surprisingly, the overall average values obtained by combining both the smooth and rough track conditions resulted in an almost perfect linear relationship.

As can be seen in Figure 4, at a data filter setting of 4,000 there is approximately a 30 percent reduction in the profile index that would be obtained with the setting at 8,000. A reduction of approximately 7 percent of the 8,000 setting value occurs for every 1,000-unit change in the data filter setting. An analysis of variance indicated that the filter setting had a

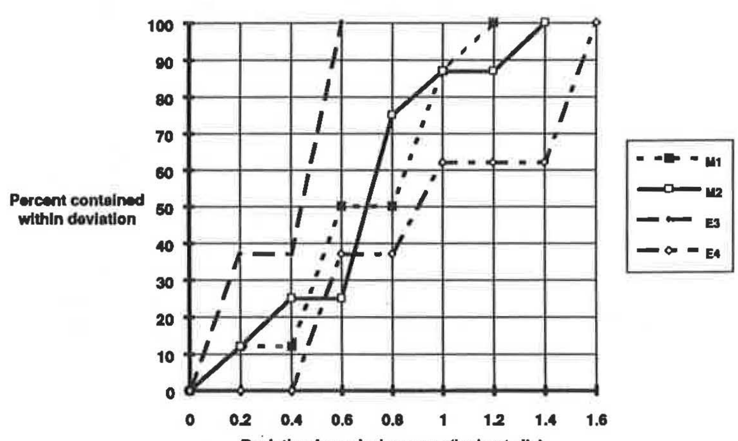

FIGURE 2 Device repeatability depicted in terms of cumulative percentage of device reading within given deviation from the device mean on rough track.

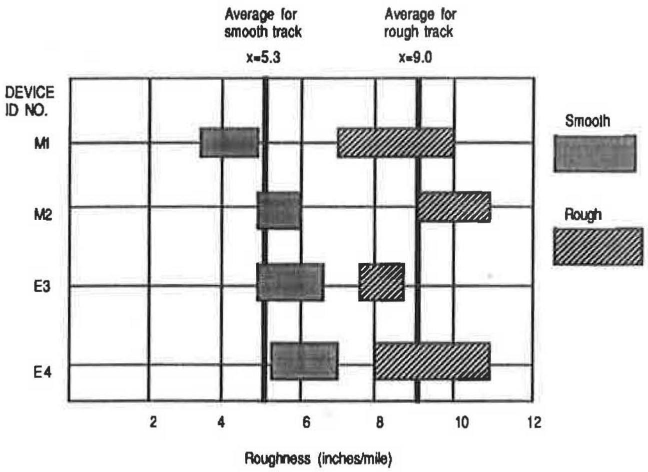

FIGURE 3 Range of roughness readings obtained during testing.

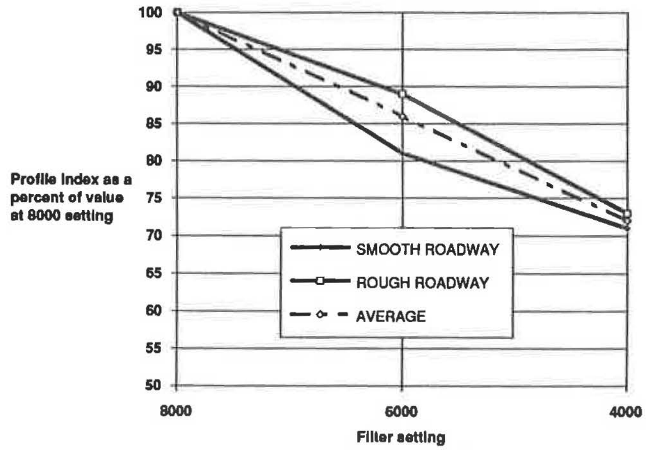

FIGURE 4 Effect of filter setting on profile index values.

significant effect on the roughness value. Duncan's multiple range test confirmed that each setting was distinct at a 1 percent significance level. The variability due to devices was significant at the 1 percent level, and the two operators differed at the 5 percent level. The analysis of variance also indicated significant interactions between operators and road roughness and between operators and devices.

The average difference between operators was 0.61 in.; the average difference between devices was 0.65 in. The difference due to filter setting was almost 1 in. between each setting level (i.e., 8,000 = 7.4 in., 6,000 = 6.4 in., 4,000 = 5.3 in.). One of the problems associated with attempting to determine specific effects from these settings is the strong interactions between operators and road roughness and between operators and devices. For example, although the operators got nearly the same average value on the smooth sections, their average values differed by more than 1 in. for the rough condition. Similarly, whereas the two operators obtained almost the same average readings on Device E4, their average readings for Device E3 differed by more than 1.5 in.

Trace Reduction Variability

Operator variability consists of field variability and trace reduction variability. Field variability is a result of the operators' inability to traverse the same path each time, measure the designated path location, and test at the same speed. It also is affected by test procedures and equipment calibration. The trace reduction variability is produced by the operator with mechanical devices. Once an operator obtains a profile trace from a mechanical device, it must be manually interpreted.

TABLE 2 PROFILE INDEX DATA OBTAINED AT DIFFERENT FILTER SETTINGS OF ELECTRONIC PROFILOGRAPHS

			Profile Index readings (in/mile)					
			Smooth wheel path			Rough wheel path		
Filter Setting	Profilograph Number	Operator	Test 1	Test 2	Mean	Test 1	Test 2	Mean
8000	E3	1	5.00	5.50	5.25	9.50	11.00	10.25
		2	4.00	4.50	4.25	8.00	8.00	8.00
	E4	1	6.00	6.00	6.00	9.50	11.50	10.50
		2	6.50	5.50	6.00	9.50	8.00	8.75
	Mean		5.38	5.38	5.38	9.13	9.63	9.38
6000	E3	1	4.50	3.50	4.00	8.00	8.50	8.25
		2	4.00	3.50	3.75	8.50	7.50	8.00
	E4	1	4.00	6.00	5.00	9.50	8.00	8.75
		2	4.50	5.00	4.75	8.00	9.00	8.50
	Mean		4.25	4.50	4.38	8.50	8.25	8.38
4000	E3	1	4.50	4.00	4.25	7.00	8.00	7.50
		2	3.50	3.00	3.25	6.00	5.50	5.75
	E4	1	3.00	3.00	3.00	8.50	6.00	7.25
		2	4.50	5.00	4.75	7.00	7.00	7.00
	Mean		3.88	3.75	3.81	7.13	6.63	6.88

TABLE 3 RESULTS OF TRACE REDUCTION BY DIFFERENT OPERATORS

		Profile index readings (in/mile) for different operators							
	Reading	Profile trace number							
Operator	Replicate	1	2	3	4	5	6	7	8
1	1	5.0	9.0	5.5	8.0	8.5	9.0	11.5	12.0
	2	3.0	5.0	3.0	6.0	7.5	8.0	10.0	11.0
	Difference*	2.0	4.0	2.5	2.0	1.0	1.0	1.5	1.0
2	1	3.5	4.5	5.0	6.5	7.0	8.0	9.5	10.5
	2	3.0	3.5	6.0	7.0	7.0	7.5	16.0	11.5
	Difference	0.5	1.0	1.0	0.5	0.0	0.5	6.5	1.0
3	1	1.5	3.0	1.5	2.5	7.0	6.5	8.0	9.0
	2	2.0	2.5	2.5	3.0	7.5	7.0	9.0	6.0
	Difference	0.5	0.5	1.0	0.5	0.5	0.5	1.0	3.0
4	1	1.5	2.5	3.0	3.0	6.5	7.5	9.5	9.5
	2	1.5	3.0	2.0	2.5	7.0	5.5	6.5	8.0
	Difference	0.0	0.5	1.0	0.5	0.5	2.0	3.0	1.5

* Absolute difference

This allows considerable judgment to be exercised in the trace analysis. An example of such a judgment factor would be whether the individual performs "outlining" before evaluating the trace.

The results of the trace reduction experiment are summarized in Table 3. All readings for the smooth wheelpath were within three standard deviations from the mean, which was 3.78 in./mi. For the rough wheelpath, one reading (16.0 in./mi) was out of control or outside the three standard deviation interval for the mean of 8.58. It should be noted that other than the extreme lack of repeatability by Operator 2 on the occasion in which the first and second readings were 9.5 and 16 in./mi, Operator 2 was the most consistent of the operators. His range was 1.0 in. for the other sets of data. The actual calculations by Operator 2 were rechecked and verified against a possible error in calculation or in recording. The extreme of 16 is worthy of concern because it shows that the present system is not adequate to prevent an out-of-control point, even by an excellent operator. A summary of the differences between first and second readings for each operator is given in Table 4.

Measurement error variability due to the difference between the operators and the repeated readings accounted for 67 percent of the product characteristic variability. There was more variability between the average values among operators than there was variability between two readings of a single operator.

Pavement Roughness Variability During Day

To evaluate changes in roughness due to time of day, a K. J. Law 690DNC profilometer was used. Before testing with the profilographs, five tests were conducted with the profilometer at 50 mph. These tests were typically conducted between 7:30 a.m. and 8:30 a.m. on each of the three test days. A second set of five profilometer tests was conducted between 10:00 a.m. and 11:00 a.m. This corresponded with the latter portions of the main profilograph experiment. This 3- to 4-hr window of testing was designed to establish whether the pavement roughness changed during the profilograph testing.

The K. J. Law 690DNC is an ASTM Class I profile measurement device. Two profile statistics were used in this testing: the Mays roughness index and the international roughness index (IRI). Mays units are expressed as inches per mile and represent the response of the vehicle to the effects of both wheelpaths. IRI units are also expressed in inches per mile. However, IRI represents the response of the vehicle to the effects of a single wheelpath. That is, the IRI unit is computed individually for the right and left wheelpaths. A total IRI can be computed by averaging the values obtained by the right and left wheelpaths.

The roughness measured by the profilometer indicated a decrease in roughness between morning and afternoon readings of 7 to 10 percent for the three test dates. The rate of decrease in roughness was 2 to 3 in./mi/hr. The Mays statistic is used for this comparison. Unfortunately, no direct comparison between the Mays units and profile index was established for this study.

During the morning testing with the profilometer, many tests included a situation termed "lost lock and saturation." This condition can be caused by excessive sunlight entering beneath the shrouds of the test van. This can result in higher-than-actual readings. The profile traces were not processed for these spikes.

An analysis of the profilograph data indicated no statistical difference between readings obtained in the morning and those obtained in the afternoon. Presumably, the large variation in profilograph test results masked the small changes in pavement roughness.

TABLE 4 DIFFERENCES BETWEEN OPERATORS' FIRST AND SECOND TRACE READINGS BASED ON FOUR TRACES FOR EACH TRACK TYPE

	Absolute differences between first and second readings (in/mile)					
Operator Number	Smooth Track			Rough Track		
	Minimum	Maximum	mean	Minimum	Maximum	mean
1	2.00	4.00	2.63	1.00	1.50	1.13
2	0.50	1.00	0.75	0.00	6.50	2.00
3	0.50	1.00	0.63	0.50	3.00	1.25
4	0.00	1.00	0.50	0.50	3.00	1.75
All operators	0.00	4.00	1.13	0.00	6.50	1.53

Change in Pavement Roughness With Time

Because the testing was conducted on 3 days over 3 months, an assessment of the change in pavement roughness with time was possible. The pavement increased in roughness by 7 percent for the morning and by 9 percent for afternoon readings. It is surprising that the morning readings had a perfectly linear relationship. The rate in change in roughness was 0.14 in./mi/day and 0.10 in./mi/day for the morning and afternoon conditions, respectively. It should be noted that this increase in roughness occurred at 5 to 8 months after construction.

DISCUSSION OF RESULTS

Historically, the use of profilographs for construction quality control has consisted of evaluating concrete pavement profiles soon after placement. The paved surfaces have typically been assessed in accordance with a maximum acceptable profile index of 7 in./mi. Recently, specifications have evolved from simple acceptance criteria to an incentive/disincentive requirement.

The precision of this device is not acceptable for the administration of incentive/disincentive specifications currently being

applied. Although some controversy currently exists regarding the correlation between mechanical and automated devices that use signal processing techniques, the actual problem is more basic than this recent development. The current specifications simply expect too much from the California profilograph.

Although a detailed statistical analysis was conducted for this study using two levels of roughness, only one section of pavement surface, 1/10 mi long, was used for all testing. This does not represent a spectrum of pavement surface types and roughness. It did, however, allow the effect of the main variables and the interaction of the variables to be clearly seen.

The industry "benchmark" standard of 7 in./mi was established before slip form paving and long before electronic paver controls. Similarly, the relationship between this numerical index and ride quality was based on the operating characteristics of vehicles from the 1950s and earlier. It is difficult to believe that modern-day pavers produce similar quality pavements and that modern vehicles respond similarly to their 1950 counterparts. A clear need exists to reexamine the industry benchmark. The evaluation should consider the quality of pavement available from modern pavers and the response to ride quality provided by modern vehicles.

The industry benchmark may well be reestablished as a function of the roadway classification or use. For example, an urban freeway with extremely high traffic volumes would appear to warrant higher standards of smoothness than rural roadways with significantly less traffic. One benchmark, for all types of roadways, does not appear appropriate for the wide range of pavement conditions found today.

Significant spatial variability exists on pavement surfaces. This variability is not easily accounted for by averaging profile traces obtained in wheelpaths. Currently, little or no information is available to determine if statistical sampling methods need to be developed to properly assess "true" roughness. Although the automated devices have significantly reduced test time, performing multiple runs under current procedures appears impractical. A study should be undertaken to determine the required sampling frequency for proper determination of representative pavement roughness values. Recent research conducted by Janoff suggests that measurement of pavement roughness ". . . can be simplified to be based on profile type roughness measured in only one wheelpath by a far simpler and less costly device than a profilometer" (*3*). This is contrary to the authors' experience.

The change in pavement roughness for both daily cycles and short-term roughness increases is not well documented. Although the phenomenon has been reported for many years, its impact on profilograph testing has not been adequately recognized. This factor should be further defined so that its implications on test timing and methods can be properly evaluated.

A surprising result from this study was the strong statistical interactions between some of the variables. These interactions make it difficult to account for the variability present in profilograph testing with limited experimentation such as with one machine or one operator.

Manual trace reduction appears to have a larger effect on the final answer than commonly believed. The average repeatability established in this study was approximately 0.94 and 1.88 in. Although not rigorously evaluated, other studies found trace interpretation repeatability to be approximately 1 in./mi (*4,5*). It is interesting to note that in one of these studies a computer-generated profilograph trace was supplied to the operator for reduction. The null blanking band was superimposed on the trace by the computer. Therefore, the null band (i.e., the template) was already depicted on the plot. The only variability measured was that of the operator's interpretation. This suggests that operator interpretation alone may approach a variability of 1 in./mi. These results strongly encourage the use of the more efficient computerized profilographs.

As incentive specifications reward contractors for producing ever smoother pavements, consideration should be given to the effects this may have on concrete mix design and resulting concrete quality. Mix designs which promote smooth pavements may produce surfaces with greater attrition and hence lower skid properties with time. Smooth pavement surfaces should be provided in concert with durable concrete pavements and not in lieu of them.

RECOMMENDATIONS

The California profilograph does not appear to have the accuracy necessary to appropriately administer a viable incentive/disincentive specification in view of the smoother and smoother pavements now possible. The industry should move away from the profile index standard and adopt some other summary statistic such as IRI or RMSVA. Using these or other acceptable profile-based statistics would require more accurate measuring equipment. They also provide a cradle-to-grave roughness statistic—that is, the statistic that would be used by the pavement designer could be directly related to the as-constructed roughness and future pavement performance.

Improvements in concrete pavement ride quality appear to have been brought about largely by the adoption of incentive/disincentive specifications and improved construction equipment. These improvements should continue to be encouraged by such specifications. However, devices for acceptance testing must be commensurate in accuracy with the monetary actions represented by these specifications. If this is not possible, specifications that set only a maximum allowable roughness level should be used.

REFERENCES

1. *Six Profilographs Readied for District Use*. California Division of Highways, May 1959.
2. D. C. Montgomery. *Introduction to Statistical Quality Control*. 2nd ed. John Wiley & Sons, Inc., New York, N.Y., 1991, Chapter 6.
3. M. S. Janoff. *NCHRP Report 308: Pavement Roughness and Rideability Field Evaluation*. TRB, National Research Council, Washington, D.C., 1988.
4. R. S. Walker and H.-T. Lin. Profilograph Correlation Study With Present Serviceability Index. In *Transportation Research Record 1196*, TRB, National Research Council, Washington, D.C., 1988, pp. 257–275.
5. M. A. Bower. *An Evaluation of Automated Profile Index Computation for James Cox and Sons, Inc. Model 8200 Profilographs*. Materials Division, Virginia Department of Transportation, Richmond, June 1991

Publication of this paper sponsored by Committee on Surface Properties–Vehicle Interaction.