**FINAL REPORT**

# QUALITY ASSURANCE OF ROAD ROUGHNESS MEASUREMENT

**Kevin K. McGhee, P.E.**
**Senior Research Scientist**



Virginia Transportation Research Council
(A Cooperative Organization Sponsored Jointly by the
Virginia Department of Transportation and
the University of Virginia)

In Cooperation with the U.S. Department of Transportation
Federal Highway Administration

Charlottesville, Virginia

May 2000
VTRC 00-R20

## ABSTRACT

This report describes a study designed to establish the accuracy and repeatability of high-speed, road-profiling equipment. This effort focused on a series of roughness validation sites. The study addresses selecting the appropriate sites, determining the "true" or reference profiles for these sites, implementing high-speed testing procedures, and developing a method for processing profiles and reducing the corresponding roughness indices.

The findings suggest that laser-equipped, high-speed inertial profilers, driven by skilled operators, are capable of providing exceptional levels of repeatability within an acceptable bias for a standard reference device (the Face Companies' Dipstick®). The study further offers some preliminary repeatability and accuracy goals for equipment used to conduct construction acceptance testing (for smoothness) as well as for network roughness surveys.

**FINAL REPORT**

# QUALITY ASSURANCE OF ROAD ROUGHNESS MEASUREMENT


**Kevin K. McGhee, P.E.**
**Senior Research Scientist**


## INTRODUCTION

### Background

The Virginia Department of Transportation (VDOT) is responsible for maintaining 90,000 centerline kilometers of interstate, primary, and secondary roads. In recent years, VDOT has resorted to contract pavement condition surveys to assess the interstate and primary road networks. To contend with the large volumes of surface, these surveys have been conducted with automated to semi-automated equipment. As recently as 1997 (the third year of the contracted condition assessments), VDOT paid $15.43 per kilometer to have 21,319 kilometers of roadway surveyed. Including the rental of special workstations necessary to view, manipulate, and report these data, three years of contracted condition surveys cost the state a total of $1.7 million.

The summer of 1999 marked the fourth construction season of an innovative program to encourage new pavement smoothness. In 1996, the maintenance overlay schedule subjected 66 kilometers of resurfacing in 2 districts to a pilot special provision for rideability. For the relatively few projects that were part of this resurfacing endeavor, the targets for smoothness were easily achieved, and the performing contractors received substantial incentives. During this past season of 1999, work that has been subjected to this special provision for smoothness has commenced on nearly 500 kilometers of new surface in 7 construction districts. Although the pavement smoothness targets have changed little, the program expansion has already resulted in projects that have not only failed to receive incentives, but also have been subject to modest disincentives. Recent trends in achieved ride quality suggest that these disincentives (and perhaps incentives) are not substantial enough to motivate contractors. If harsher disincentives become necessary, the special provision will most certainly invite increased scrutiny from the industry.

A common element to network condition assessment and the special provision for smoothness is the measure of pavement ride quality. When measuring pavement condition, ride quality is an important indicator of overall serviceability. For the new construction special provision, ride quality is the only issue of significance.

## Equipment History

Traditionally, testing pavement ride quality has involved devices that are referred to as response type road roughness meters (RTRRMs). The estimates of roughness obtained from these devices are based on the response of an instrumented vehicle (or trailer) traveling at a designated speed over a pavement surface. Unfortunately, the data produced by RTRRMs tend to be very instrument-specific and subject to wide variability. This is because any change that affects the response of the vehicle to the surface (e.g., suspension or tire wear, vehicle and/or operator weight changes, etc.) will also affect the estimate of roughness. In order to achieve any semblance of repeatability (either from one instrument or from a fleet of instruments), the calibration requirements must be rigorous. Typically, these calibration procedures involve operating the devices on roads that have been previously accepted as reference surfaces and making the necessary adjustments to the hardware to ensure that the output is consistent with the established reference measurements. The obvious difficulties with this approach are that the "true" roughness of these surfaces are difficult to determine, and once determined, these true values are subject to change over time and from season to season.[1]

In recent years, industry has replaced many of the traditional roughness measuring instruments that indirectly measure longitudinal profiles with equipment that directly measures the longitudinal profiles. One very common example is an instrument known as the South Dakota type Road Profiler (SDRP). SDRPs are more technically referred to as high-speed inertial road profilers. SDRPs conform to the American Society for Testing and Materials (ASTM) Standard E-950. Inertial profiling systems use a combination of accelerometers, height sensors, and electronic distance measuring instruments to collect road profiles. Calibration of these systems involves calibration of the individual components. With the exception of the distance-measuring equipment, these components can be calibrated without moving the host vehicle. The integrated distance-measuring component is calibrated like any other vehicle-mounted, distance-measuring instrument.

Most importantly, the profile-based approach used by inertial road profilers theoretically eliminates any influence from the host vehicle.

## The International Roughness Index

The fundamental purpose of a properly functioning inertial road profiler is to collect accurate longitudinal profiles. Profiles as an end product, however, have little practical use in most instances. It is more common to generate statistics that summarize the character of these profiles, particularly those characteristics that contribute to road roughness. Perhaps the most widely used statistic used to summarize profile character is the International Roughness Index (IRI). The IRI is produced using a sophisticated quarter-vehicle model and a measured longitudinal profile. This reference quarter-car model is complete with all of the basic parameters necessary to describe an actual automobile (or at least a critical portion of it). These parameters include: mass of the vehicle body, suspension, wheels and tires; spring stiffness coefficients for the vehicle springs, shocks and/or struts; and damping coefficients indicative of a conventional shock-absorbing system. The simulated suspension motion is accumulated and divided by the distance traveled to yield the IRI.[2] Smaller values represent a smoother ride and higher values are indicative of a rougher one.

In Virginia, the IRI is used to assess the ride quality of most newly constructed pavement (including maintenance resurfacing) and all of the existing surfaces tested as part of the annual network level surveys.

### Calibration/Validation

Calibration, in its traditional sense, is not the same issue with inertial profilers that it has been with RTRRMs. For the most part, it is simply a matter of following the manufacturer's directions for maintaining the calibration of the individual components. However, inertial profilers sometimes produce erroneous data. The sophistication of these devices introduces a new series of potential problems. Minor malfunctions can completely debilitate a profiler. Worse yet, elusive circuit shorts and modest changes in settings can produce subtle errors in data that are not easily resolved or even noticed (until it is too late).

What this may suggest is that the critical issue with profilers is not so much calibration as it is validation. A quality verification process would not entail tuning output through hardware adjustments. An inertial profiler is either functioning properly or it is "invalid." To ensure that inertial profilers are functioning such that the resulting summary indices are accurate and reliable, the most practical contemporary answer is a regular validation program.

## PURPOSE AND SCOPE

This report describes a study designed to establish the accuracy and repeatability of high-speed, road-profiling equipment. This study had two objectives. The first was to provide a mechanism to ensure the quality of the roughness data collected for pavement management (which is typically done by contract). The second was to establish a plan to verify the quality of data collected with agency-owned equipment. A common and fundamental concern to both issues is reliability, or more specifically, measurement validation. The corresponding basis of this study is a formal plan to establish a measure of the accuracy and repeatability of roughness equipment, particularly for equipment that is owned by VDOT. In other states, as well as in Virginia, it is anticipated that the performance of DOT-owned equipment could then be used to establish reasonable precision and bias requirements for contracting equipment.

This formal plan presented in this study involved testing a series of roughness validation sites. Sites were selected to represent each of the three major highway systems, the most common road surfaces, and a full range of riding quality.

# METHODS

Establishing the roughness validation program involved selecting a pool of test sections, determining the “true” profiles of these test sections, establishing a standard test procedure and schedule for high-speed profiling, conducting the tests, processing the resulting profiles, and generating the corresponding roughness indices. The resulting summary roughness indices were compiled to estimate the accuracy and repeatability of the high-speed equipment and its operator.

## Validation Site Selection and Preparation

The sites were selected to provide typical levels of ride quality on common Virginia pavement surfaces. The sites were distributed to include most regions of the state. The sections were chosen to avoid extremes in geometry (curves and hills) or changes in surface type and texture. To promote safer lane closure (as necessary), good sight distance was also a prerequisite. As often as possible, validation sites were selected to coincide with the Strategic Highway Research Program’s Long-Term Pavement Performance (SHRP LTPP) sites. This alignment with the SHRP program was beneficial because those sites were already striped in such a way as to facilitate easy identification and testing. If needed, the LTPP databases also contain a comprehensive performance history on these sites, as well as independent measures of ride quality. When the validation sites were not SHRP sites, every effort was made to lay them out consistently with the SHRP methods.

The schematics in Figures 1, 2, and 3 describe the layout of a typical validation site. Marking a site involves establishing a 305-meter “run-up” section with a contiguous 150-meter “profiling” section and approximately 75 meters of “run-down” pavement. The *profiling* section is obviously the most important. The *run-up* section is required when high-speed tests are conducted under lane-closure conditions. An accurately established *run-up* distance is also helpful when the inertial equipment does not include a photo-triggering device. The *run-down* interval ensures that there is an adequate overlap profile among the respective instruments and between tests. This *run-down* interval also reminds the high-speed equipment operator to maintain constant speed throughout the entire *profiling* section.

Although a surveyor’s tape is recommended, a large (1.2-meter circumference) measuring wheel was used to establish the locations of each of the marks. A fully “marked” validation site contains four 0.5 meter paint stripes extending from the right edge of the pavement into the test lane. Once the test sections were selected and the general limits established, the wheel-paths were located. The exact position of the wheel-paths relative to the centerline and to the edge of the pavement depended on the width of the lane. Regardless of the lane width, however, all wheel-paths were located 1.75 meters apart to correspond with the known spacing of the outside sensors of the high-speed profilers. While initially measuring the 150-meter profiling section, additional tick marks were placed at 8.2-meter intervals. Using the centerline or edge striping as a reference, the wheel-paths were marked at each of the 8.2-meter intervals. Next, a crew of three people used the wheel-path indicators to place chalk-lines longitudinally for the extent of the 150-meter test section. Lastly, transverse chalk-lines were placed at the beginning and end of the profiling section and 300-mm radius semi-circles were drawn around each corner of the resulting box.

## Reference Profiles

With the respective wheel-paths established and clearly marked, the next step was to collect the “true” profiles. The accepted reference for measuring true profiles is the precision rod and level method (described in ASTM Method E1364). However, for this and most modern calibration and comparison studies,[3,4,5,6] a device known as the Dipstick® was used to provide a mechanized substitute.

The Dipstick®, manufactured by the Face Companies, is generically referred to as a static inclinometer. It is a manually operated device that accumulates elevation data along a line in precisely controlled steps. These accumulated elevations combine to produce very accurate profiles. Figure 1 offers an illustration of the “boxing” technique used to enhance the accuracy of the Dipstick®-based profiles. This technique consists of a continuous survey down one wheel-path (A), across the lane at the site end (B), a return walk down the opposite wheel-path (C), and then closing the box back to the starting point (D). For more information on collecting a Dipstick® profile, the reader is urged to review the proceedings of the 1993 Road Profiler User Group meeting.[7]

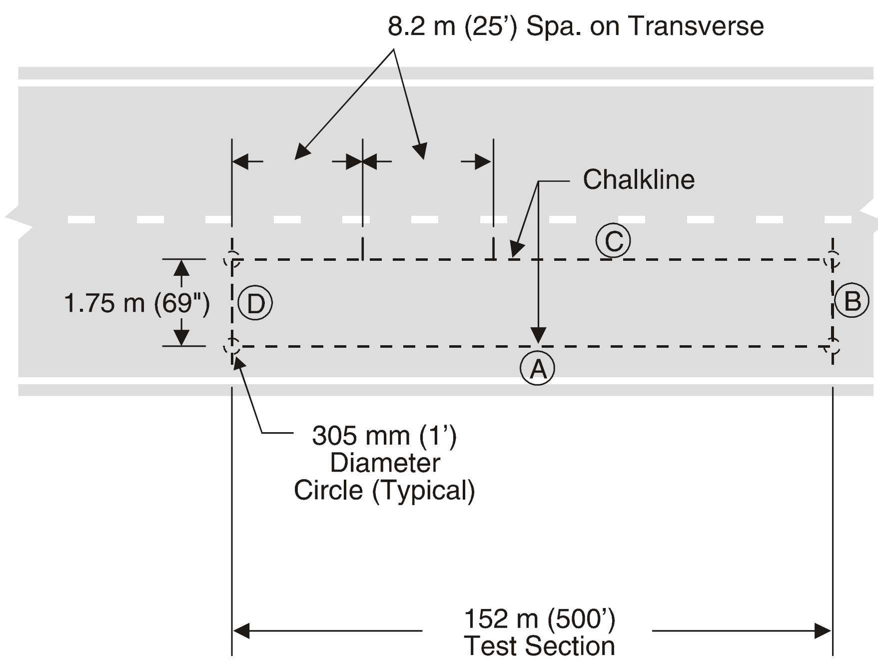


**Figure 1. Site Layout for Reference Profiling**

## High-Speed Test Procedures

The next step in the validation process involved collecting profiles with the high-speed inertial equipment. Figure 2 describes the layout of an entire validation site, including the minimum run-up and run-down sections that were necessary for high-speed tests. Although it is impossible with modern technology to completely separate the operator and the instrument, this series of full-speed tests was designed to focus on the equipment. Generally, these tests involved the first 5 to 7 runs performed by the high-speed equipment immediately upon completion of the survey conducted with the Dipstick®. The tests were run under traffic control and with the benefit of either artificial bumps or an optical trigger mechanism.

Figure 2 also illustrates the use of artificial bumps. In essence, their presence inserts a spike in the resulting profile at known distances from each end of the test section. By locating the spike in the data, an analyst can precisely extract only that portion of the profile that corresponds with the reference profiles. Incidentally, early tests with artificial bumps indicated that locating them too close to the beginning of a test section produces a falsely high roughness reading in the subsequent subsection (this phenomenon is discussed more thoroughly in a previous report).[8] The initial artificial bump should therefore be located at least 8 meters in advance of the test section being tested.

Near the end of the initial round of validation testing, an optical trigger mechanism was added to the profiler. For testing with the optical trigger, the artificial bump was replaced with a 0.5-meter section of highly reflective tape (Figure 3). With the automatic trigger, it was only necessary to begin the tests a sufficient distance back (or a minimum of 100 meters) from the reflective tape in order to achieve full test speed with an adequate run-up "waste" profile. The optical trigger inserts an automatic event code into the profile data. The reporting software responds to these codes by creating section breaks.

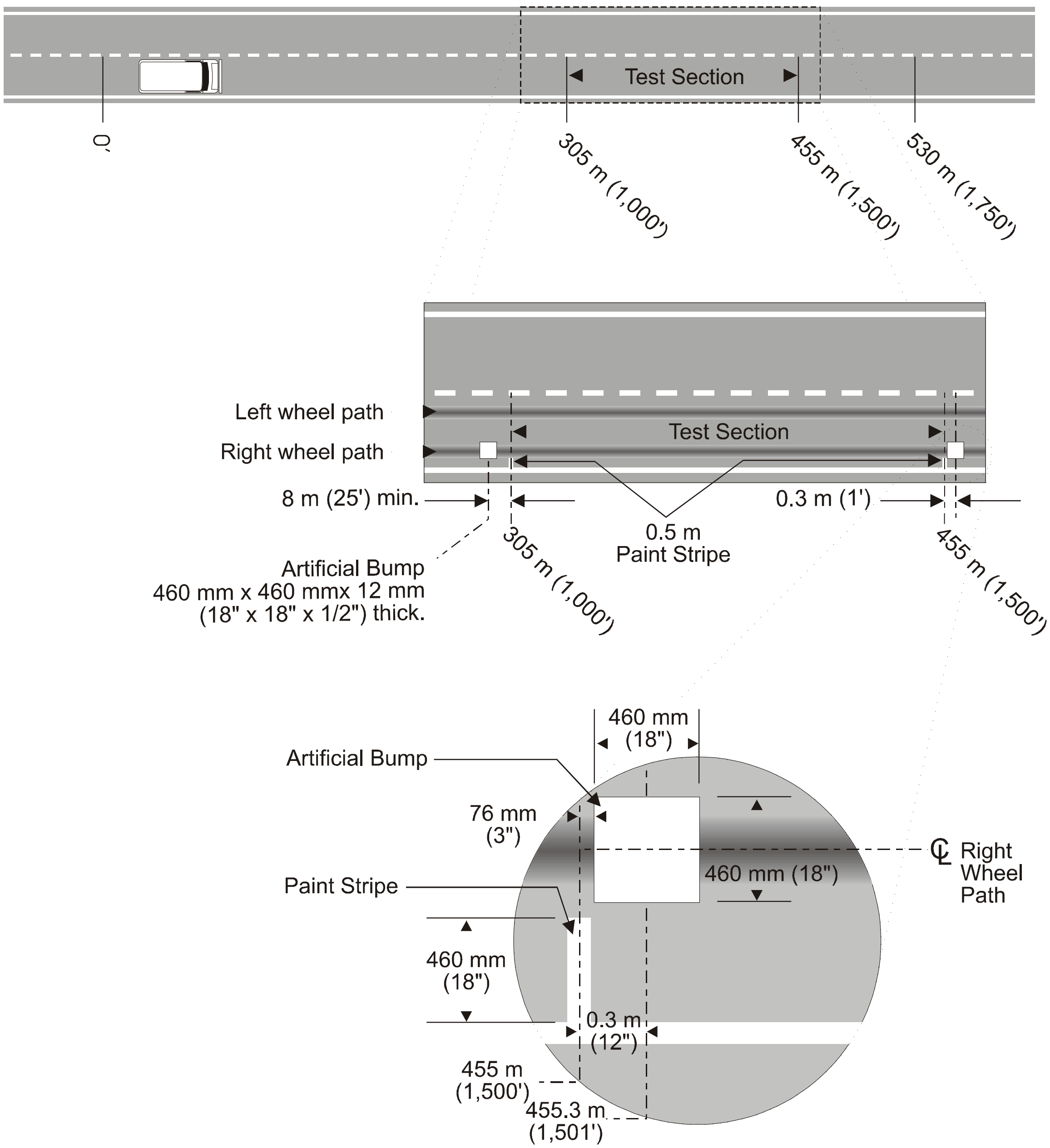


**Figure 2. Test Site with Artificial Bumps**

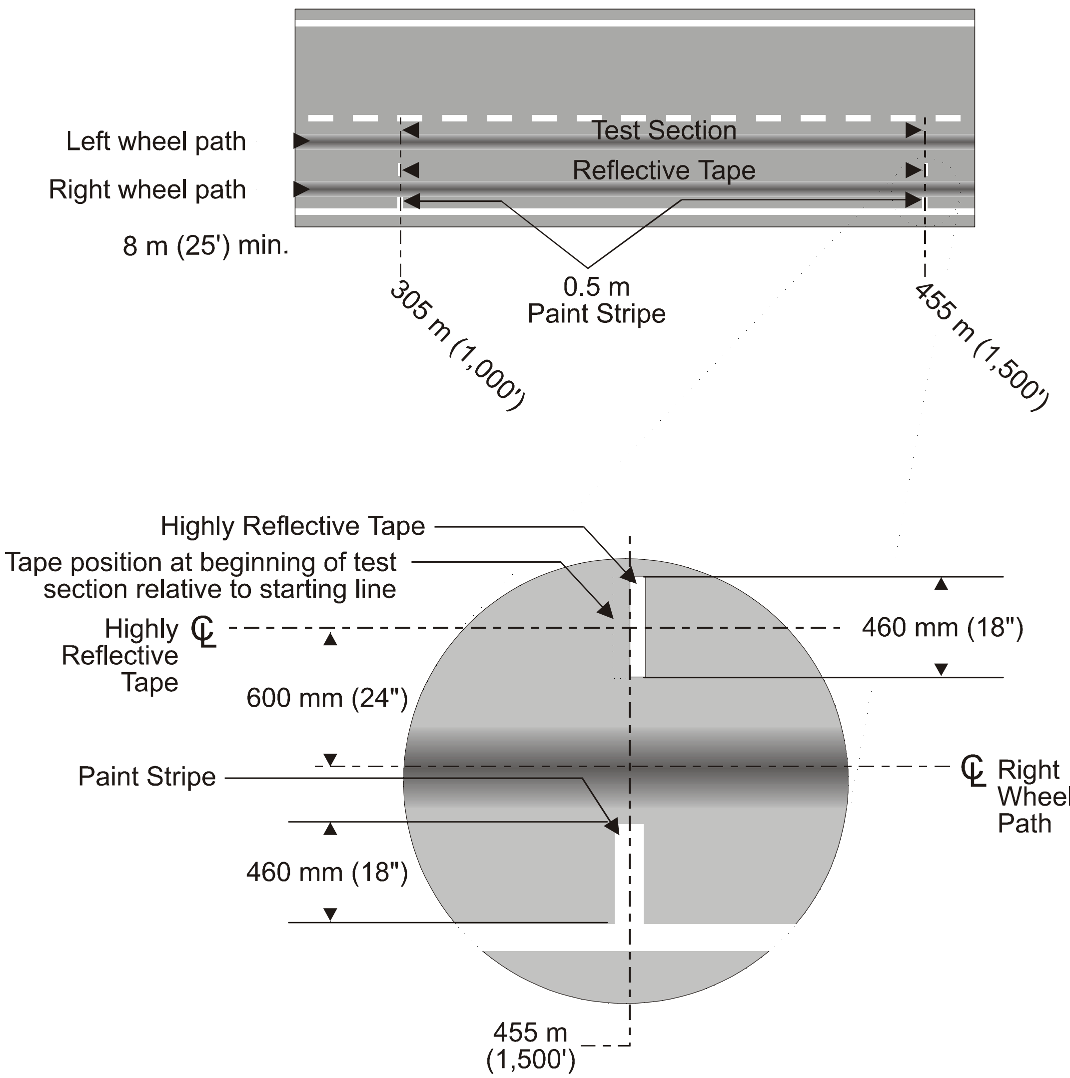


**Figure 3. Test Site with Photo-Triggering Switch**

## Reducing Profiles

Elevation profiles from both the static inclinometer (Dipstick®) and the inertial profiler were imported into a third-party software package designed to easily generate a variety of statistics relevant to highway profiles. Before these profiles could be imported, however, data gathered from these two types of equipment required some post-processing within the respective fabricator's reporting software.

### Face Companies' Dipstick®

The first step in processing a Dipstick® data set, descriptively referred to as "unboxing," involved correcting and disassembling the continuous profile into its four individual legs. The "correcting" that was performed on the unboxed profiles involves forcing the beginning and end of the continuous profile to align vertically, and then distributing the accumulated error evenly over the profile's entire length. The next step involved creating ASCII-type elevation files from the two rectified wheel-path profiles (box legs A and C from Figure 1). Finally, a special descriptive header was inserted into each ASCII file and the data were made available to the analysis program.

### Inertial Profilers

Profiles originating from high-speed equipment ultimately conformed to the same requirements as those originating from the Dipstick®. The reporting software provided by the instrument fabricator was used to assemble profiles from raw sensor data. Typically, profiles for both wheel-paths were generated for the entire length of the test, including the run-up and run-down sections. Next, the portion of the longer profile that corresponds with the profile test section was extracted. When artificial bumps were in place (which is one method of identifying test section limits), these limits were identified by locating the spikes in the elevation data. When the optical trigger switch was available (which is a second method of identifying test section limits), the section limits were identified automatically. When neither method was practical (e.g., when runs were conducted without traffic control), the beginning and end of the test section had to be identified by the operator during the run. Obviously, this manual approach was less precise. Fortunately, there was sufficient "vertical character" otherwise present within most profile sections to match the test limits with previously established limits simply through graphical comparisons.

From this point, the data from the inertial profilers were treated identically to the data from the Dipstick®. The ASCII elevation files were fitted with appropriate header information and readied for more intense analysis.

### UMTRI Software

The profile comparison analysis was performed using RoadRuf, an integrated set of software tools distributed through the University of Michigan's Transportation Research Institute (UMTRI). In addition to providing a more exotic profile interpretation, RoadRuf also offers a well-tested mechanism for computing common roughness indices. For the purpose of this study, the RoadRuf software served as a consistent, objective tool to generate IRI values for each wheel-path of each test run for each device. These values were then assembled into a result matrix for further analysis.

## Repeatability and Accuracy

The initial result matrix was assembled with the intention of answering two questions: How successfully does this equipment reproduce results over consecutive tests for the same roadway surface? And, how well does Virginia's high-speed profiling equipment compare to reference equipment?

### Repeatability

The first issue concerns the precision that is possible within the limits of a single operator/instrument combination. More technically, this situation is referred to as repeatability under repeatability conditions, with repeatability conditions described as the same operator using the same instrument within a short period of time (i.e., within an hour). A method for addressing repeatability of profilers using IRI measurements is offered by Karamihas, et al., in the *Guidelines for Longitudinal Pavement Profile Measurement*.[9] In this reference, individual IRI values are normalized by the average IRI obtained from multiple runs. These normalized values are then accumulated in a histogram. The amount of scatter in this histogram provides an indication of the instrument's repeatability. Karamihas and his colleagues acknowledge that the scatter (and thus the repeatability) could be represented by the standard deviation for the distribution. However, they further propose that a more relevant way to evaluate a number of instruments (a fleet of profilers or competing contract vendors) is to establish limits on scatter and to assess how many individual measurements fall within those limits. Their study suggests that the number of values falling within 2 percent of the average provides a measure of how well an instrument would serve as a *reference* device (a *reference* against which other high-speed profilers or RTRRMs may be compared). Furthermore, the number of values that fall within 5 percent provides a measure of a profiler's suitability for network-level survey work.

The current study used the above-described methods to examine repeatability of a single operator/instrument combination. This study also uses the "Karamihas" methods to not only evaluate the suitability of this combination for network and reference work, but also to serve as a construction acceptance tool.

### Accuracy

The second important issue is accuracy (or bias). This analysis addressed accuracy by comparing the IRI values generated from profiles produced by a reference device with those originating from the high-speed equipment. Like the repeatability analysis, the accuracy approach involved "normalizing" the individual IRI measurements. In this case, the individual measurements were normalized against the IRI measurements generated using the Dipstick®. For example, if the Dipstick®-based IRI was 1000 mm/km and a profiler-based IRI for the same wheelpath was 900 mm/km; the normalized IRI would be 900 divided by 1000 or 0.90. As an additional step, these normalized values were reduced by 1.0 and multiplied by 100 to obtain a percent bias from the reference. The bias of the measurement in this example would therefore be

–10 percent. The average of all of these values provided a bias for the respective instrument/operator combination.

Accuracy is an elusive concept for high-speed profiling equipment, primarily because a "ground-truth" profile and roughness index is difficult to establish. The Dipstick® is the most commonly used reference device for modern road-profiling equipment. However, there are physical differences between inertial profilers and the Dipstick® that may be responsible for legitimate and somewhat expected disagreements. Figure 4 provides a footprint comparison of the two devices. The beam from a profile-grade laser is approximately 2 mm in diameter, while the Dipstick® typically uses a circular pad that is more than 30 times larger (64-mm diameter). Furthermore, the inertial equipment uses averages that are taken periodically from nearly continuous sampling to represent profile elevation points (see the thicker vertical lines in Figure 4). The Dipstick®, on the other hand, accumulates elevation profiles purely through discrete steps. In one sense, it seems likely that the minute relative footprint and high sampling rates associated with laser-equipped profilers may make these profilers more sensitive to smaller features. On the other hand, the averaging technique used with laser data may provide a *smoothing* effect to the profile that is not realized with the Dipstick®-based data.

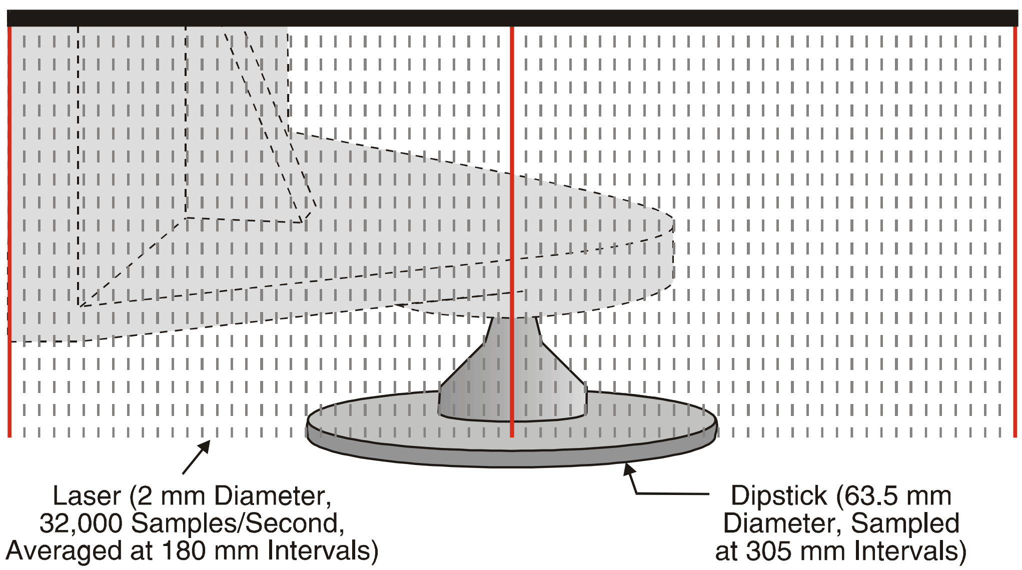


**Figure 4. Sensor Footprint Comparison**

Perhaps the most obvious difference in the profiles produced by the two devices is evident when longer features are of interest. To help explain, Figure 5 is an illustration of two unfiltered (raw) profiles for the same wheelpath on the same test section. The lighter line represents the profile as "seen" through the Dipstick®. The heavier line is an inertial profiler's "perspective" of the same profile. Observe that both profiles are able to distinguish the joint deterioration present at approximately 10-meter intervals along the test section. Also notice, however, the expected insensitivity of the inertial profiler to a design grade of just under one percent.

Figure 6 depicts the identical profiles after a 30-meter *high-pass* filter has been applied. That is, all features (wavelengths) longer than 30 meters have been removed, allowing shorter wave (higher frequency) features to *pass*. In this illustration, the Dipstick® continues to respond more literally to the extremes of the profile (higher "highs," and lower "lows"). However, the general character of the two profiles is remarkably similar.

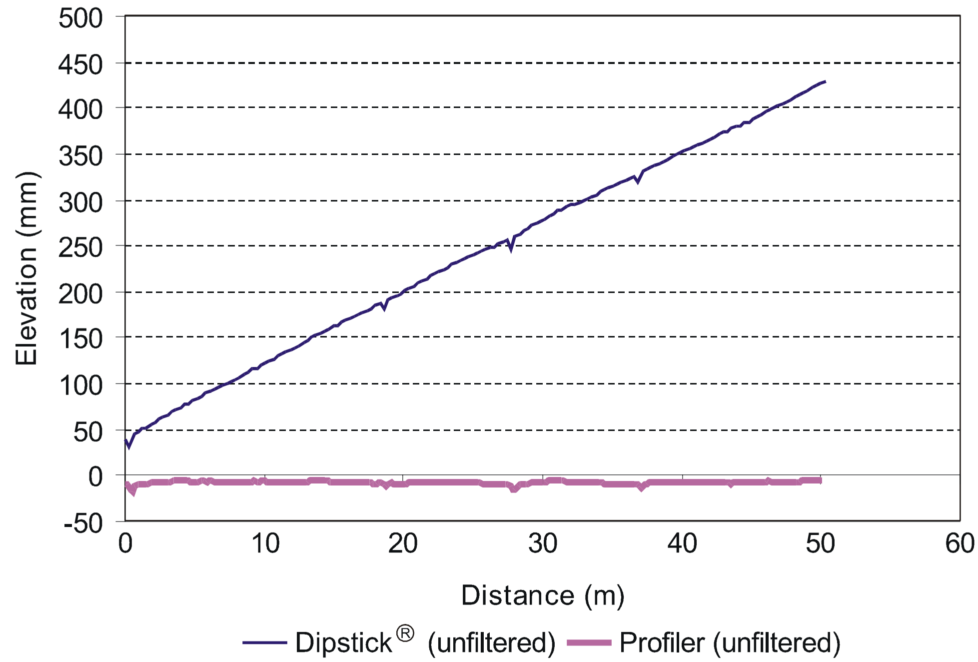


**Figure 5. Unfiltered Profiles**

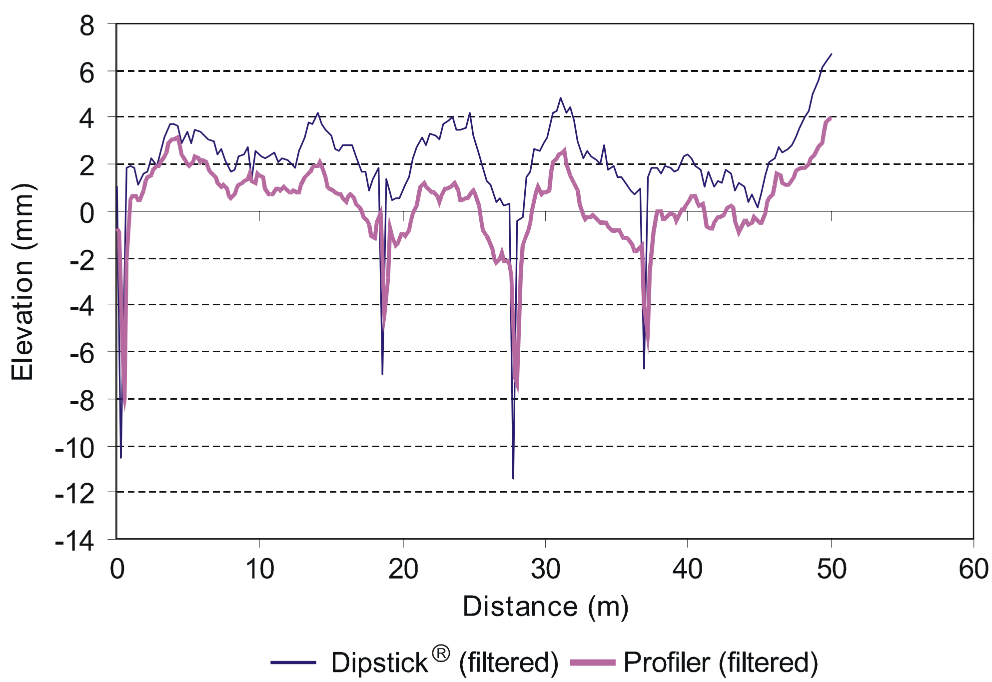


**Figure 6. Filtered Profiles**

While percent bias may be a good perspective from which to evaluate the overall accuracy of profilers, it may not be the best method to approach disagreement between the reference device and the profiler on a case-by-case basis. There is no reason to expect the additional roughness “seen” by one profiler (and not another) to be proportional to the overall roughness associated with a longitudinal profile. For example, with a Dipstick®-based profile with an estimated 2000 mm/km IRI, long vertical curves can conceivably contribute 200 mm/km of additional IRI roughness, or 10 percent. This same 200 mm/km may not show up in an inertial profiler’s data. Conversely, consider the situation where the “design” profile has exactly the same component of long vertical curvature, while the rest of the profile contributes only another 800-mm/km IRI roughness. In this instance, the proportion of additional roughness identified by the Dipstick®, but not the inertial profiler, is 20 percent. Theoretically, the same problem could arise at the other end of the wavelength spectrum, except the instrument that may fail to see the increased roughness would be the Dipstick®, and not the inertial profiler.

Fortunately, the IRI algorithm attenuates long (greater than 30 meters) and short (less than 1 meter) wavelengths and tends to focus on features within a waveband that both instruments are capable of measuring accurately. However, the theoretical likelihood of systematic error prompted an absolute bias assessment for several sites where the disagreements were more acute.

## FINDINGS AND DISCUSSION

By spring of 1999, 12 validation sites had been selected and profiled using the reference device (Dipstick®), as well as an inertial profiler operated by the Research Council. Geographically, there are three sites in the Suffolk District, one in Bristol, one in Salem, two in Richmond, two in Fredericksburg, two in Lynchburg, and one in Staunton (see Figure 7). Pavement types included surface treatments, conventional hot-mix asphalt, continuously reinforced concrete, and composite pavements. The smoothness (or roughness) of the calibration sites ranges from an IRI of 630 mm/km to more than 4700 mm/km. In order to include samples of extremely rough pavement, two of the test sites were located on secondary roads. The remaining categories of ride quality (smooth to moderately rough) were represented by 1 interstate site and 9 divided primary sites. Lane widths varied from approximately 3 to 4.25 meters.

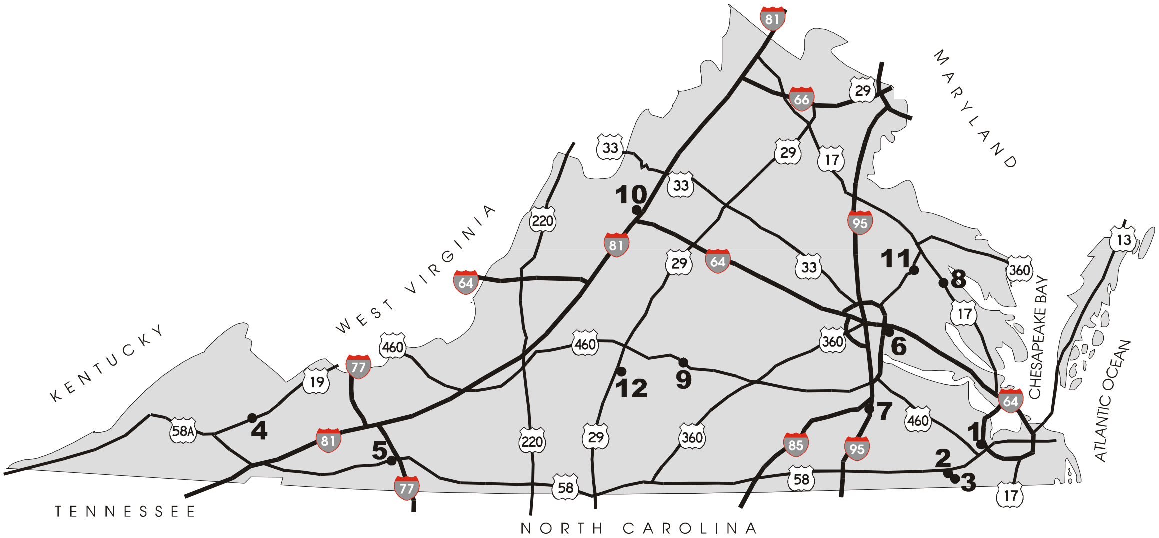


**Figure 7. Validation Site Map**

## Repeatability Expectations

Two profiles were produced for each test run, one each for the left and right wheel-paths. A single summary IRI was generated for each profile. The product of all the tests was 12 sites x 2 wheelpaths x 5 repeat runs = 120 IRI values. Figure 8 provides the distribution of normalized IRI values for an entire data set. Again, each IRI is expressed as a ratio of the average for the respective site. That is, if an individual test run produced an IRI of 980 mm/km while the average for that wheel-path was 1000 mm/km, the normalized IRI for that test would be 0.98.

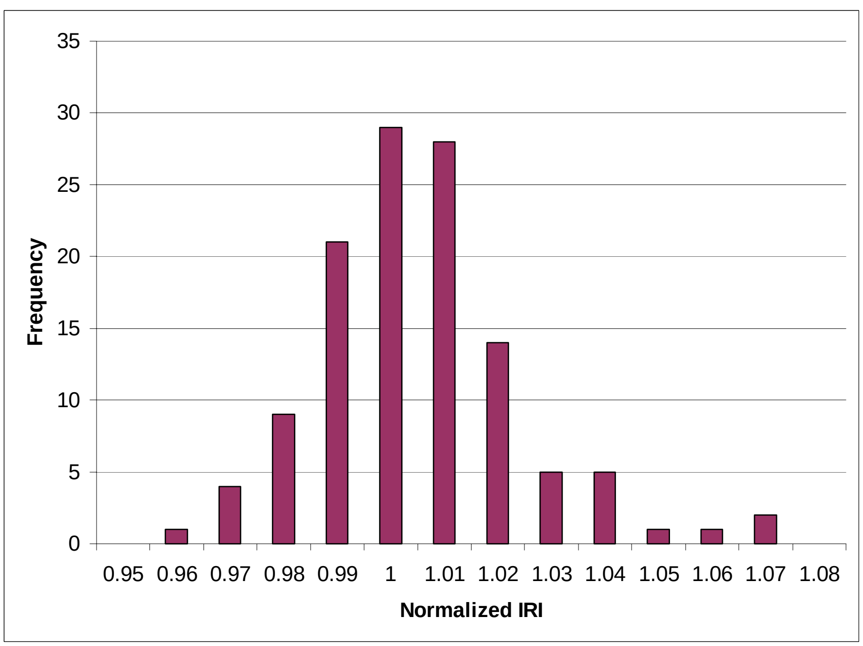


**Figure 8. Repeatability of Profiler**

The observed repeatability of this particular profiler/operator combination was excellent. For 120 tests, there were no individual measurements below 0.96 and none above 1.066. The standard deviation of this distribution is less than .02 (2 percent).

Table 1 provides the results of the analysis designed to observe the performance of the profiler/operator within the proposed scatter limits of 2 and 5 percent. At this point, it is helpful to share some results of a similar repeatability analysis conducted using data from the 1993 RPUG experiment. In that experiment, 33 different devices were tested on 30 sites in 4 regions of the country. Karamihas' analysis of this data found only 2 device/operator combinations that were able to achieve within 2 percent repeatability better than 75 percent of the time.[9] Similarly, only two devices were able to exceed 5 percent repeatability more than 98 percent of the time. This comparison confirms the exceptional repeatability observed for the combination of the profiler and operator that participated in this study.

**Table 1. Repeatability of VDOT Profiler/Operator**

| No. of Tests | Std. Dev. (%) | Within 2% | | Within 5% | |
|---|---|---|---|---|---|
| | | (Count) | (%) | (Count) | (%) |
| 120 | 1.89 | 92 | 76.7 | 117 | 97.5 |

## Accuracy Expectations

Figure 9 is the distribution of bias percentages produced by the accuracy analysis for this profiler/operator combination.

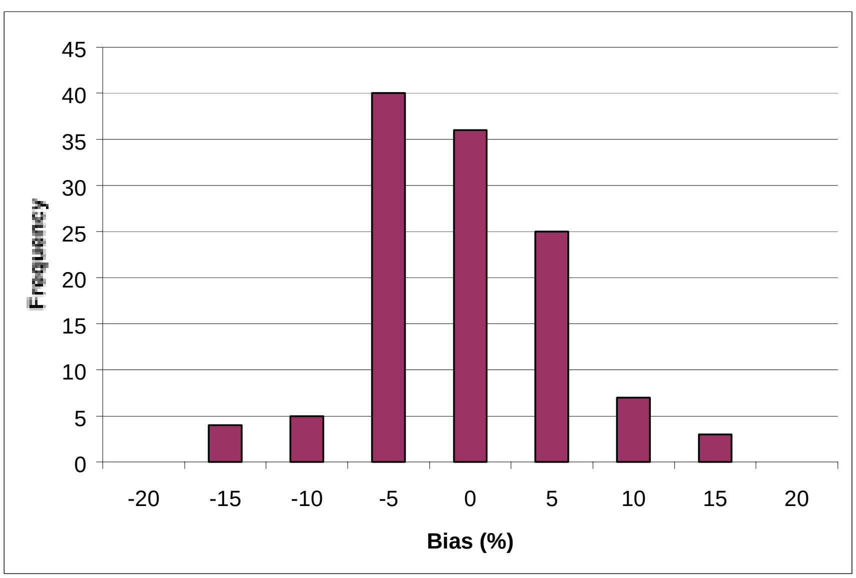


**Figure 9. Bias in IRI**

The average of this distribution is –2.89, which means that this particular profiler and operator combination underestimates the IRI by an average of just under 3 percent. Table 2 summarizes the performance of the profiler/operator within an arbitrary accuracy limit of 5 percent. In the 1993 RPUG experiment, 27 profilers were examined. In that analysis, the participating profilers overestimated the reference IRI by an average of 22 percent. Of the 27 profilers in the RPUG study, 10 were optical or laser-based systems. The average bias for those 10 systems was an improved +7.93 percent. Still, only two profilers in the 1993 study exhibited an absolute bias of 2.9 percent or less. Likewise, these two systems were the only ones to achieve this overall accuracy by producing IRI values within 5 percent of the reference more than 50 percent of the time.

**Table 2. Accuracy of VDOT Profiler/Operator**

| No. of Tests | Bias (%) | Within 5% | |
|---|---|---|---|
| | | (Count) | (%) |
| 120 | -2.89 | 61 | 50.8 |

The bias in Figure 9 exhibited a remarkable clustering around zero. It is clear, however, that there were individual readings that differed considerably from the reference. Furthermore, many of these larger differences originated from a select number of validation sites. Table 3 contains the average percent bias, by wheel-path, for the three sites on which the disagreements were most conspicuous and consistent.

**Table 3. Extremes in Bias**

| | Left Wheel-path | Right Wheel-path |
|---|---|---|
| Site | Bias (%) | Bias (%) |
| 4 | -6 | -6 |
| 9 | -8 | -12 |
| 12 | -16 | +10 |

Further scrutiny into the reason for these higher disagreements yielded some interesting, if not counter-intuitive results. To explain, first consider Figure 10, which illustrates elevation profiles from a segment of one of the validation sites (site 4). The predominant character of these profiles is described by a "textbook" sinusoidal wave of approximately 25 meters in length. Although subtle high-frequency (short wavelength) differences exist, the overall character of the two profiles are remarkable similar and the IRI's produced are in nearly perfect agreement (829 and 828 mm/km for the Dipstick® and SDRP, respectively).

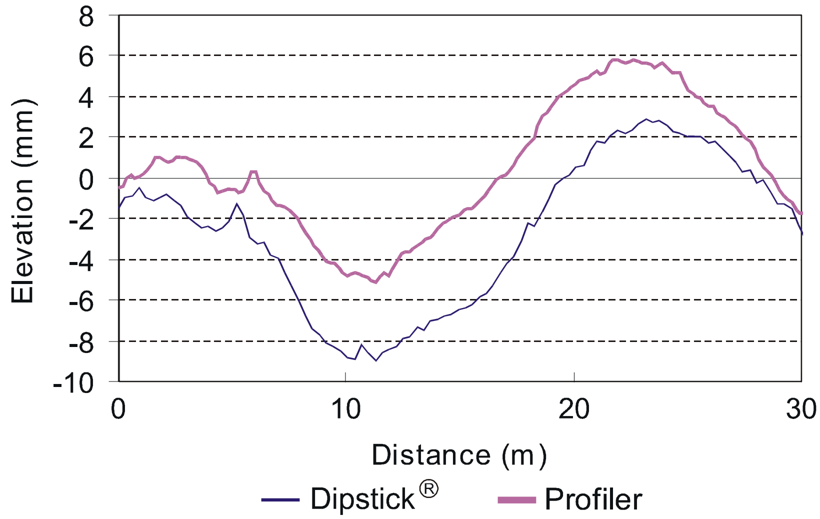


**Figure 10. Profiles with Agreeing IRI Values**

Figure 11 depicts profiles from the next 30-meter segment of the same validation site. The predominant character is much different. This segment's roughness appears to be produced by higher-frequency (shorter-wave) features. Unlike the previous segment, there was a measurable disagreement between the IRIs produced from the Dipstick® and SDRP-based profiles. In fact, the SDRP-based IRI (1281 mm/km) under-predicted the Dipstick®-based IRI (1486 mm/km) by 14 percent.

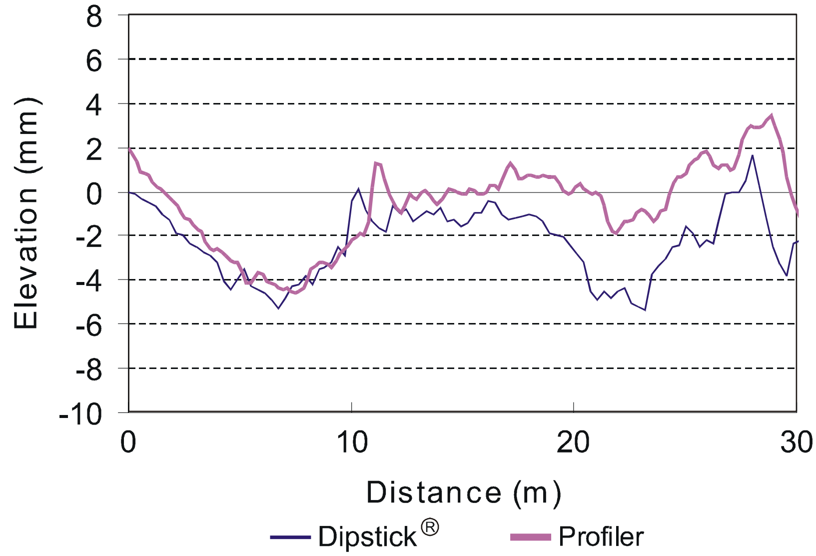


**Figure 11. Profiles with Disagreeing IRI Values**

The tendency for the SDRP-based IRI values to underestimate the reference IRI was observed repeatedly, especially among profiles (or portions of profiles) that were constructed of higher-frequency (shorter-wave) roughness. The one notable exception was the right wheel-path of site 12. In this instance, the added resolution available to the SDRP-based profiles appeared to reveal some short-wave activity that was not detected by the Dipstick®.

# CONCLUSIONS

## Repeatability & Accuracy

High-speed profilers serve two primary functions in Virginia. The first is to provide the roughness component of network distress surveys. The second function is using these profilers as a construction acceptance tool relating to special provisions for ride quality. With these functions in mind, the following suggestions are offered to aid in the establishment of repeatability and accuracy standards:

### Repeatability

*Construction Acceptance*

This study has demonstrated that a high-speed profiler and operator combination can expect to estimate an IRI within 2 percent of the average (from many runs) in at least 3 out of 4 test runs (or 75 percent of the time). To understand how this level of expected repeatability impacts the equipment's ability to administer a construction provision, it is helpful to know more about the specific testing procedures. In Virginia, each lane of a new overlay is subjected to a minimum of two repeat profiler runs (more runs are used if the first two tests are not in reasonable agreement). The pay adjustments are calculated using the data from the test run that produces the lowest summary roughness index. An overlay project involving two lanes, for example, would require a minimum of 4 repeat profiler runs. Of these 4 runs, only 1 would be expected to return values outside of the 2 percent limit on scatter for repeatability. Unless this "bad" value were inordinately low (which would be an advantage to the contractor), it would almost certainly be inconsequential.

*Network Surveys*

For network survey work, the IRI estimate on any one pavement segment is rarely of critical importance. Consequently, the precision requirements (in terms of scatter limits) for equipment commissioned to collect high-volume roughness data can be relaxed slightly. However, since repeat runs are rarely practical, and the opportunity to identify and correct 'bad" tests is almost nonexistent, the need for consistency may be more important. Although a slightly larger scatter limit may be acceptable, recorded values should remain within this limit a high percentage of the time. The ability of a profiler/operator combination to estimate an IRI within 5 percent of the average and 95 percent of the time, is consistent with these expectations. The findings of this study suggest that this is certainly an achievable level of repeatability.

### Accuracy

This research acknowledges that systematic bias can and does exist. However, the reference pavement sites used in this study were carefully selected to represent the range of surfaces commonly encountered on Virginia highways. Similarly, the equipment conforms to the standards expected for all road roughness measurement conducted for VDOT. The bias that was observed in this study is therefore considered reasonable for any equipment that is operated by or for the State of Virginia. Accepting some additional error associated with slightly less control over test conditions, an average bias of not more than 5 percent (as compared to the Dipstick® device), is recommended when these 12 sites are used. In general, profilers should be able to achieve within 5 percent of the reference value 50 percent of the time when one is unable to use all or any of these sites.

## Formal Validation Process

In order to determine the repeatability and accuracy requirements of a given vehicle/operator combination, it is recommended that the process demonstrated through this report be conducted on a regular basis. The series of validation sites established through this study are offered for use by operators of VDOT equipment. These processes and these sites are also recommended for contract equipment and operators, although some subset of the total validation program may be acceptable. This does not mean that the proposed standards should be relaxed; only that they may be established on fewer and less geographically diverse sites.

In summary, the findings of this study support the following conclusions:

- Laser-equipped, high-speed inertial profilers, driven by skilled operators, are capable of producing exceptional levels of repeatability.
- Laser-equipped, high-speed inertial profilers, driven by skilled operators on a variety of conventional highway surfaces, are capable of reproducing the roughness estimates of a standard reference device within a 3 percent bias.
- Contrary to expectations, the Dipstick® was repeatedly observed to be more sensitive to higher-frequency roughness than a laser-based inertial road profiler.

## RECOMMENDATIONS FOR FURTHER RESEARCH

Considering the complexity of equipment and diversity of surfaces encountered in this effort, the ability of the high-speed inertial profilers to provide the same results in consecutive runs is quite remarkable. It is significant that during all of these tests, the high-speed profiler was driven by the same operator on the same day and with very little time between each repeat run. An extremely important element, particularly when the concept of contracted roughness surveys is considered, is the reproducibility between different equipment and operators.

To further confound things, it is known that road profiles (and correspondingly ride quality) change over time and from season to season. Understanding seasonal influences is also important for managing pavements, particularly with respect to determining the optimum times to conduct yearly distress surveys.

To address these issues, it is recommended that two additional studies be conducted. The first study would build on the implementation of the validation measures proposed in this report. It would focus on the repeatability and accuracy characteristics of changing equipment and operators. Within a very few cycles (years) of the formal validation program, an excellent opportunity will exist to assess the reproducibility of road roughness measurement.

The second recommended study would concentrate on time and seasonal effects on road profile measurement. One component of this second study would use the cyclic nature of the validation program to observe the natural changes in measured profile from year to year. A shorter-term component should be designed to study early-age (days, weeks, and months) changes in profile due primarily to curing, compaction, and exposure to traffic.

## REFERENCES


1 Gillespie, T.D., Sayers, W.M. and Segel, L. 1980. *Calibration of Response-Type Road Roughness Measuring Systems. National Cooperative Highway Research Program*. NCHRP Report No. 228. Washington, DC: Transportation Research Board.

2 Sayers, M.W. 1996. *On the Calculation of International Roughness Index from Longitudinal Road Profiles*. *Report No. TRR 1501.* Washington, DC: Transportation Research Board.

3 El Korchi, T. and Collura, J. 1998. *A Comparative Study of Ride Quality Measuring Devices*. Report No. TRR 1643. Washington, DC: Transportation Research Board, pp. 125-135.

4 Perera, R.W., Kohn, S.D., and Bemanian, S. 1996. *Comparison of Road Profilers*. Report No. TRR 1536. Washington, DC: Transportation Research Board, pp. 117-124.

5 Sayers, M.W., Karamihas, S.M. 1996. *Estimation of Rideability by Analyzing Longitudinal Road Profile*. Report No. TRR 1536. Washington, DC: Transportation Research Board, pp. 110-116.

6 National Research Council. *Manual for Profile Measurement: Operational and Field Guidelines* 1994. Strategic Highway Research Program. Washington, DC.

7 Perera, Rohan W., and. Kohn, Starr D. 1994. *Road Profiler Data Analysis and Correlation–Final Report*. Research Report No. 92-30. Plymouth, Michigan: Soil and Materials Engineers, Inc.

8 McGhee, K.K. 1999. *Measuring, Achieving, and Promoting Smoothness of Virginia's Asphalt Overlays*. VTRC Report No. 19. Charlottesville, Virginia: Virginia Transportation Research Council.

9 Karamihas, S.M., et. al. 1999. *Guidelines for Longitudinal Pavement Profile Measurement*: Final Report. NCHRP 10-47. Washington, DC: Transportation Research Board.